

# **Modelação e implementação de um sistema de *Business Intelligence* para tratamento de *Big Data* agrícola**

**Caso de estudo Herdade Vale da Rosa**

**Pompeu Mendes Jorge Pais Dias**

Dissertação para a obtenção do Grau de Mestre em  
**Engenharia Agronómica**

Orientadores: Doutor Luís Manuel Bignolas Mira da Silva

Licenciado Bruno Ricardo Pancadas Pereira Caldeira

## **Júri:**

Presidente: Doutora Elisabete Tavares Lacerda de Figueiredo Oliveira, Professora Auxiliar do(a) Instituto Superior de Agronomia da Universidade de Lisboa.

Vogais: Doutor Luís Manuel Bignolas Mira da Silva, Professor Associado do(a) Instituto Superior de Agronomia da Universidade de Lisboa, orientador;

Doutor Ricardo Nuno da Fonseca Garcia Pereira Braga, Professor Auxiliar do(a) Instituto Superior de Agronomia da Universidade de Lisboa.

## Agradecimentos

Ao meu orientador Professor Luís Mira da Silva pelo incentivo e motivação, sempre valorizando o trabalho realizado e a realizar, pela excelente orientação em momentos de definição da dissertação e, em especial, pela constante disponibilidade que permitiu o progresso nas várias etapas do trabalho.

Ao meu orientador Bruno Caldeira, por me introduzir ao tema do *Business Intelligence* em agricultura, por todo o acompanhamento durante a fase de aprendizagem e também durante o trabalho prático realizado, pela partilha de conhecimentos fundamentais para a resolução de problemas e pela sua simpatia e disponibilidade.

Aos membros da Herdade Vale da Rosa, em especial ao Engenheiro Joaquim Praxedes, à Doutora Sara Domingos e ao Pedro Ramalho, por permitirem a realização deste trabalho, pelo espírito crítico e disponibilidade fundamentais para o desenvolvimento desta ferramenta e pela hospitalidade com que me receberam nas diversas reuniões que realizámos.

Agradeço a todos os Professores do Instituto Superior de Agronomia, que, de diferentes formas, despertaram o meu interesse pela Engenharia Agronómica e pelos desafios que a mesma enfrenta, o que me levou a desenvolver um trabalho diferente, com base num problema real e atual no setor agrícola.

A todos os meus amigos e colegas que me acompanharam durante todo o meu percurso académico, com um agradecimento especial para Guilherme Paulo, Luís Rebelo de Andrade, Henrique Carvalho, Pedro Moreira da Silva, Luís Graciosa e Joana Nuncio, por me providenciarem a força para acreditar que iria concluir o meu mestrado com sucesso, encorajando-me incessantemente durante estes últimos exigentes meses.

À minha família pelo seu apoio incondicional para sempre atingir o que mais ambiciono, mas, principalmente à minha mãe, que sempre me deu as condições, força e motivação para todos os dias ser uma pessoa melhor e, nesta última etapa académica, pela incessante disponibilidade para me ajudar e pelo interesse no meu trabalho.

## Resumo

A análise de dados em agricultura representa hoje um importante desafio em todo o processo produtivo. A variedade, velocidade e volume com que os dados são produzidos em algumas empresas fazem com que os métodos tradicionais de processamento de dados não sejam suficientes para uma tomada de decisão ótima. Torna-se assim pertinente aliar os conhecimentos agronómicos, analíticos e informáticos para colmatar esta limitação através de sistemas que realizem automaticamente a transformação de dados em informação útil sobre a produção agrícola, num curto período. De facto, os temas de digitalização e *Big Data* surgem como fundamentais para a produção agrícola lograr uma intensificação sustentável, de forma a ultrapassar os exigentes desafios de competitividade económica, segurança alimentar e ambientais.

Este trabalho visou a criação, implementação e validação de um sistema de *Business Intelligence* para processamento e exposição de informação proveniente de *Big Data* agrícola, de forma a melhorar o processo de tomada de decisão.

Foi realizado um caso de estudo na empresa Herdade Vale da Rosa, onde existe um substancial investimento tecnológico conducente a um elevado volume e variedade de dados relativos aos processos agrícolas. Estes dados encontram-se em constante atualização, em função das diferentes atividades realizadas, e são armazenados numa base de dados.

Neste trabalho, foi efetuada uma primeira obtenção de dados e criado um modelo de demonstração para a apresentação do potencial do sistema. Posteriormente, foi realizado um levantamento das necessidades de informação dos utilizadores para criação e implementação de modelos ajustados às necessidades dos mesmos. Os modelos finais criados foram avaliados e validados por diferentes utilizadores.

Em suma, o processamento de dados realizado pelo sistema desenvolvido, representa uma oportunidade para uma agricultura mais eficiente na utilização de recursos, para uma maior qualidade de produção e para a otimização da produtividade, ou seja, uma oportunidade para a intensificação sustentável agrícola.

**Palavras-Chave:** Digitalização; *Big Data*; *Business Intelligence*; Agricultura de precisão; Apoio à tomada de decisão.

## **Abstract**

Data analysis is considered today one of the most prominent challenges in the agricultural production process. The velocity, variety, and volume in which data is produced in some agriculture businesses makes traditional processing methods obsolete for an optimal decision-making process. It is thus crucial to merge agronomic, analytical and computer knowledge to overcome this limitation through the creation of automatic systems that can import and process agricultural data on a short period of time. In the agricultural sector, digitalization and Big Data are already regarded as fundamental to achieve a sustainable intensification of the agricultural production, and by achieving so, overcome even bigger challenges as world food security and environmental safety.

This work aimed to create, implement, and validate a Business Intelligence system for processing and exhibiting information originated from agricultural Big Data, thus improving its users' decision-making process.

A case study was developed with Herdade Vale da Rosa, an agricultural business with a strong emphasis on technological investment leading to a vast volume of agricultural data, with substantial variability and in constant update.

Firstly, in this case study, a data collection corresponding to different areas of activity was carried out, along with a first demonstration model that was presented to the users. Then, an inquiry was done to perceive the user's information necessities for posterior modeling of the program's final version. Lastly, a final program evaluation was executed through another inquiry with the users. This allowed validating whether the program was suited for the business information needs.

In conclusion, the data processing performed by the developed system constitutes an agriculture opportunity for a more efficient input usage, for a better quality production, as well as a productivity optimization, in other words, an opportunity for agricultural sustainable intensification.

**Key-words:** Digitalization; Big Data; Business Intelligence; Precision agriculture; Decision support system.

## Índice

Agradecimentos.....	i
Resumo .....	ii
Abstract.....	iii
Lista de abreviaturas .....	vii
Índice de quadros .....	viii
Índice de figuras .....	ix
1. Introdução.....	1
1.1. Agricultura de precisão e <i>Big Data</i> agrícola.....	1
1.1.1. Evolução da agricultura de precisão .....	2
1.1.2. Intensificação sustentável e agricultura de precisão.....	3
1.1.3. Fatores limitantes em agricultura de precisão .....	5
1.1.4. <i>Big Data</i> .....	6
1.1.5. <i>Big Data</i> agrícola .....	7
1.1.6. Agricultura de precisão como precursora para a utilização de <i>Big Data</i> agrícola .....	8
1.1.7. Importância e oportunidades agrícolas na utilização de <i>Big Data</i> .....	9
1.1.8. Desafios do <i>Big Data</i> no setor primário.....	10
1.2. Sistemas de apoio à tomada de decisão (DSS) .....	11
1.2.1. Tomada de decisão .....	11
1.2.2. Processo de tomada de decisão .....	11
1.2.3. Programas MCDM.....	12
1.2.4. Utilidade e funcionamento dos DSS.....	14
1.2.5. Diretrizes futuras dos DSS's .....	15
1.2.6. Internet como ferramenta .....	16
1.2.7. Programas <i>Business Intelligence</i> no tratamento de <i>Big Data</i> .....	18
2. Caso de estudo .....	19
2.1. Enquadramento .....	19
2.2. Objetivo geral .....	21
2.3. Objetivos específicos.....	21
2.3.1. Análise de dados .....	21
2.3.2. Implementação de modelos .....	21
2.3.3. Verificação.....	21

3. Material e Métodos .....	23
3.1. Power BI® .....	23
3.1.1. Obtenção e edição de consultas .....	24
3.1.2. Cálculos e relações entre consultas.....	25
3.1.2.1. Operações .....	26
3.1.2.2. Criação de relações.....	26
3.1.2.3. Colunas calculadas .....	27
3.1.2.4. Medidas .....	27
3.1.3. Relatórios dinâmicos do modelo ( <i>Dashboards</i> ) .....	27
3.1.3.1. Filtragem por segmentação de dados.....	28
3.1.3.2. Visuais .....	28
3.1.3.3. Partilha .....	29
3.2. Método aplicado ao caso de estudo.....	29
3.3. Dados cedidos pela Herdade Vale da Rosa .....	30
3.3.1. Dados de consumos.....	31
3.3.1.1. Registos de consumos .....	31
3.3.1.2. Programa inicial de consumos .....	32
3.3.2. Dados de operações .....	34
3.3.2.1. Registo de operações.....	34
3.3.2.2. Objetivos de operações.....	36
3.3.3. Dados de colheita.....	36
3.3.3.1. Registo de colheita .....	37
3.3.3.2. Objetivos de colheita .....	38
3.3.4. Dados culturais.....	39
3.3.5. Dados físicos .....	40
3.3.6. Dados climáticos .....	41
3.3.7. Dados de qualidade.....	42
4. Resultados da criação de modelos de BI.....	43
4.1. Modelo de demonstração .....	43
4.1.1. Planeamento do modelo de demonstração .....	44
4.1.2. Modelação .....	45
4.1.3. Modelo de demonstração .....	49
4.2. Determinação das necessidades de informação .....	54

4.2.1 Diferenciação de modelos .....	55
4.2.2. Necessidades técnicas – Consumos na fertirrega .....	55
4.2.3. Necessidades na análise de operações.....	56
4.4. Modelação final .....	57
4.4.1. Modelo de consumos .....	57
4.4.2. Modelo de operações.....	60
4.5. Modelos finais.....	65
4.5.1. Modelo de consumos .....	65
4.5.2. Modelo de operações.....	69
5. Discussão dos resultados .....	75
5.1. Avaliação qualitativa dos modelos .....	75
5.1.1. Funcionamento do sistema segundo Shim et al. (2002) .....	75
5.1.2. Desempenho em agricultura de precisão segundo McBratney et al. (2005) .....	76
5.1.3. Processamento de <i>Big Data</i> agrícola segundo Wolfert (2017) .....	77
5.2. Análise dos resultados da modelação e implementação do sistema .....	78
5.3. Perspetivas .....	79
6. Considerações finais .....	81
7. Referências bibliográficas .....	83
Anexos.....	87

## **Lista de abreviaturas**

AP – Agricultura de precisão

BI – Business Intelligence

DSS – Decision Support System

FAO - Organização das Nações Unidas para Alimentação e Agricultura

IOT – Internet of Things

PBI – Power BI®



## **Índice de quadros**

Tabela 1 - Comparação entre bases de dados convencionais e bases de dados de Big Data.....	7
Tabela 2 – Exemplo de tipologia de formatação uniforme.....	25

## Índice de figuras

Figura 1 – Ciclo de gestão <i>Smart Farming</i> alicerçado numa base de dados online e num DSS .....	16
Figura 2 – Diagrama da área de exploração da Herdade Vale da Rosa. ....	19
Figura 3 – Modelo conceptual do método aplicado na criação, implementação e validação da utilização do programa Power BI®.....	30
Figura 4 – Excerto da tabela de consumos de 2015 a 2018.....	31
Figura 5 – Excerto do programa inicial de consumos desde o abrolhamento até à frutificação. ....	33
Figura 6 – Excerto do programa inicial de consumos desde o engrossamento até à maturação. ....	33
Figura 7 - Excerto da tabela de registo de operações de 2015 a 2017. ....	35
Figura 8 – Exemplo da tabela de coeficientes de objetivos de operações para a campanha 2018.....	36
Figura 9 - Exemplo de tabela de registo da colheita da campanha 2017.....	37
Figura 10 – Excerto da tabela de previsão de produção para a campanha 2017-2018.....	38
Figura 11 – Exemplo de tabela de data de ocorrência de estados fenológicos em diferentes variedades de videira, na campanha de 2017. ....	39
Figura 12 - Exemplo da tabela original com as características físicas das quadras. ....	40
Figura 13 - Exemplo da tabela dos dados climáticos estação meteorológica Herdade Vale da Rosa. ....	41
Figura 14 - Exemplo da tabela (reduzida) do controlo de qualidade. ....	42
Figura 15 - Amostra da tabela auxiliar de calendário.....	45
Figura 16 - Mapa de relações do modelo de demonstração.....	46
Figura 17 - Painel de comparação entre a previsão de produção e a produção para o ano de 2017. ....	50
Figura 18 - Painel de consumo de água de rega para diferentes variedades, de 2014 até 2018. ....	51
Figura 19 - Painel de análise de supervisores e a sua respetiva colheita durante uma campanha.....	52
Figura 20 - Painel de registo histórico de consumos .....	53
Figura 21 - Painel de registo histórico de operações.....	54
Figura 22 - Excerto das tabelas auxiliares de designação de produtos e de dados físicos .....	58
Figura 23 - Mapa de relações do modelo de consumos.....	58
Figura 24- Excerto da tabela auxiliar "Número de trabalhadores". ....	61
Figura 25 - Excerto da tabela auxiliar "Tarefa".....	62
Figura 26 - Mapa de relações do modelo de operações.....	62
Figura 27 - Painel de registo histórico de consumos. ....	65
Figura 28 - Painel de plano inicial de consumos para 2017. ....	66
Figura 29 - Painel de execução de consumos. ....	67
Figura 30 - Painel de diferença entre execução e plano inicial de consumos. ....	68
Figura 31 - Painel de registo histórico de operações.....	69
Figura 32 - Painel de registo histórico dinâmico de operações. ....	70
Figura 33 - Painel de objetivos das operações. ....	71
Figura 34 - Painel de execução de operações.....	72
Figura 35 - Painel de diferença entre objetivos e execução de operações. ....	73

## 1. Introdução

Com o acentuado desenvolvimento tecnológico na atualidade, juntamente com a redução significativa de preços da tecnologia agrícola nas últimas décadas, a quantidade e variedade de dados para o agricultor tomar uma decisão torna o armazenamento, processamento e análise dos mesmos, limitante. Esta informação torna-se tão mais abundante quanto mais acentuada é a presença tecnológica capaz de registrar e transmitir dados (Lindblom et al., 2017). Perante esta tipologia de dados produzidos, os meios convencionais de tratamento de dados não representam uma solução adequada para os processos de tomada de decisão em agricultura, pelo que se torna fundamental o desenvolvimento de ferramentas que logrem, por um lado processar a totalidade dos dados disponíveis e, por outro, expor a informação de uma forma inteligível para o utilizador (Wolfert, 2017).

Nesta dissertação é exposto um caso de estudo em que o principal objetivo é a criação e aplicação de um programa com as características enunciadas numa empresa agrícola. Para além deste objetivo é também pretendido com este trabalho realizar uma análise de dados para criação de indicadores relevantes, implementar o programa na empresa e, por fim, verificar o desempenho do programa em condições reais de tomada de decisão através de uma avaliação realizada pelos seus utilizadores.

Este trabalho seguiu uma metodologia em 6 fases principais, respetivamente (1) a obtenção de dados, (2) criação de um modelo de demonstração, (3) determinação das necessidades de informação dos utilizadores, (4) criação dos modelos finais, (5) verificação e, finalmente, (6) avaliação e discussão de resultados.

Em concreto, com este trabalho são pretendidos dois tipos de resultados distintos. Por um lado, a criação de modelos funcionais para as áreas de operações e de consumos de fertilizantes, o que significa que no final deste trabalho se pretende que estes modelos se encontrem em condições de continuar o processamento de dados autonomamente gerando relatórios dinâmicos, sem ser necessária uma intervenção posterior. Por outro lado, a verificação de que o sistema e os modelos implementados são realmente úteis para o processo de tomada de decisão dos utilizadores. A junção destes dois resultados permitirá uma análise crítica quanto à utilização de sistemas de *Business Intelligence* em condições agrícolas reais.

### 1.1. Agricultura de precisão e *Big Data* agrícola

A evolução recente da prática agrícola conduziu à transição de uma agricultura dirigida pelas características e necessidades solo (Robert, 1993), com recurso a tecnologias de taxas de aplicação variável de produtos e, posteriormente, tecnologia de condução assistida, para uma agricultura baseada na qualidade do produto e na gestão ambiental (McBratney et al., 2005).

O setor agrícola utiliza, diretamente e indiretamente, grandes quantidades de energia na forma de máquinas, combustível, sementes, fertilizantes, estrume, químicos, eletricidade e água. O uso

excessivo de energia e a diminuição de fontes de obtenção da mesma estão entre os principais motivos para a necessidade de otimização da utilização de recursos e para alcançar uma agricultura sustentável (Erdal et al., 2007). De acordo com a Sociedade Americana de Agronomia (1989), “a agricultura sustentável é aquela que, no longo prazo, melhora a qualidade ambiental e a base de recursos de que depende; satisfaz as necessidades básicas de alimentação humana e de fibra; é economicamente viável; e melhora a qualidade de vida dos agricultores e da sociedade como um todo”.

Se admitirmos que, inevitavelmente, as explorações agrícolas necessitam de insumos externos para serem, pela definição, sustentáveis, a posse de conhecimento específico sobre o local de exploração auxilia a utilização desses recursos de forma ambientalmente responsável. A agricultura de precisão representa assim uma oportunidade para se atingir a agricultura sustentável (Bongiovanni e Lowenberg-DeBoer, 2014).

Existem várias definições de agricultura de precisão, formais e informais. Todavia o consenso parece residir em considerar a agricultura de precisão como uma forma de prática agrícola que aumenta as decisões corretas tomadas por unidade de área e por unidade de tempo, à qual estão associados benefícios líquidos. Esta definição afasta assim a agricultura de precisão da obrigatoriedade de elevada presença de tecnologia na exploração e de ferramentas de resolução espacial, pois embora as decisões possam ser baseadas em sistemas tecnológicos, como sistemas de GPS, entre outros, podem também ter por base a gestão técnica humana (McBratney et al., 2005).

Os meios utilizados para obter decisões corretas a diferentes escalas temporais e espaciais condicionam outras definições de agricultura de precisão, uma vez que para obter informações relevantes no momento de tomada de decisão em reduzidas unidades de área e de tempo, geralmente é necessário um forte investimento em equipamentos, pelo que se define também a agricultura de precisão pela utilização de técnicas e sensores geoespaciais para a identificação de variações no campo e atenuar as mesmas com estratégias alternativas (Zhang e Kovacs, 2012).

A agricultura de precisão na atualidade é também definida pela capacidade de digitalização (Lindblom et al., 2017). A digitalização define-se pela capacidade de utilização e processamento dos dados digitados, ou seja, dos dados registados num formato digital, de forma a ser acrescentado valor aos mesmos (Gray e Rumpe, 2015). No fundo, a agricultura de precisão, para ser funcional e lograr a integração de todas as fontes de dados, necessita de abordar primeiramente a digitalização de sensores, mapas, registos, entre outros, para modelos, algoritmos, medidas e visualizações de forma a utilizar eficientemente o valor acrescentado conferido pelos dados produzidos (Lindblom et al., 2017; Wolfert, 2017).

### **1.1.1. Evolução da agricultura de precisão**

Segundo Crookston (2006), a agricultura de precisão é uma das dez revoluções mais importantes que afetaram a agricultura e os ecossistemas agrícolas nos últimos cinquenta anos. Mais concretamente, a incorporação de tecnologias de GPS e de deteção remota permitiu aos agricultores uma maior eficiência na utilização de insumos agrícolas e regular as externalidades associadas.

Os primeiros registos da prática de agricultura de precisão datam dos finais da década de 1980, com a introdução ao público do sistema de GPS. Surgiu assim uma nova era tecnológica agrícola, uma vez que passou a ser possível mapear e categorizar as parcelas agrícolas a uma escala da ordem dos metros ou mesmo inferior, anteriormente de obtenção muito dispendiosa. A visão homogénea de uma parcela transformou-se numa visão heterogénea mais aproximada da realidade das características edafoclimáticas e culturais presentes nessa mesma parcela. Com esta ferramenta, as práticas de fertilização foram as primeiras a ser alteradas com o intuito de aumentar a eficiência da aplicação e, simultaneamente, aumentar a produtividade cultural. Durante os anos seguintes foi essa a principal utilização da agricultura de precisão. No entanto, as aplicações de agricultura de precisão à proteção de culturas, sistemas de rega, sementeiras/plantações e monitorização espacial de produtividade, inicialmente menos exploradas, foram também implementadas ao longo do século XX (Auernhammer, 2001).

Atendendo a que a agricultura de precisão possibilita a criação de benefícios na eficiência do uso de insumos para a produção e, por definição, tem por objetivo aumentar o benefício líquido das produções agrícolas, em Portugal, as técnicas e tecnologias de agricultura de precisão começaram por ser utilizadas em culturas arvenses no início da década de 90. Num passado mais recente, a prática de agricultura de precisão começou a ser utilizada em culturas que conferem uma maior margem bruta de retorno, de forma a amortizar o investimento tecnológico necessário, e que estão associadas a uma grande dependência de fatores de qualidade para a sua valorização comercial, como foi o caso das culturas hortícolas (Braga, 2009).

### **1.1.2. Intensificação sustentável e agricultura de precisão**

A generalização da prática de agricultura de precisão surge numa época em que existe um desafio global emergente, a obtenção de uma maior quantidade de produto com um menor nível de insumos e externalidades, designado por *intensificação sustentável*. Este desafio assenta em quatro premissas principais: o aumento da produção, o aumento da produtividade, o foco na segurança alimentar e a forma de aplicação da intensificação sustentável (Garnett et al., 2013).

O aumento da produção é um fator determinante na intensificação sustentável, uma vez que a sustentabilidade assenta não só nas necessidades ambientais, mas também nas necessidades económicas e de alimentação mundial.

O aumento da produtividade é uma premissa obrigatória neste desafio. A conversão da atual superfície agrícola não utilizada na produção de alimentos, em superfície agrícola utilizada, não é uma possibilidade eficaz. Esta conversão comprometeria a sustentabilidade ambiental, uma vez que representaria uma diminuição na capacidade mundial de captura e armazenamento de carbono (por exemplo, na forma de áreas florestais), levando a elevados custos no combate a emissões de gases com efeito de estufa (Stern, 2008). Isto significa que a produção agrícola necessita aumentar, mas ocupando sensivelmente a mesma área que ocupa na atualidade.

Segundo a FAO (1996), “existe segurança alimentar quando as pessoas têm, a todo o momento, acesso físico e económico a alimentos seguros, nutritivos e suficientes para satisfazer as suas necessidades dietéticas e preferências alimentares, a fim de levarem uma vida ativa e sã”, ou seja, a produção de alimentos necessita, não só de existir em quantidade suficiente, mas também com diversidade suficiente de forma a satisfazer as preferências e necessidades de cada indivíduo. Com o crescimento da população mundial, o tema de segurança alimentar tem, cada vez mais, um forte impacto na forma como a agricultura é praticada. Nos países em desenvolvimento, esta carência de segurança alimentar está bastante presente nas sociedades, pelo que a introdução de novas tecnologias no seu setor primário – habitualmente o setor mais importante economicamente em países em desenvolvimento – seria uma forma de alcançar uma maior produção e diversidade de alimentos, aumentando a eficiência da mão-de-obra e dos insumos (Feder et al., 1985). A prática de agricultura de precisão como forma de gestão, ao ser implementada nestes países, poderia constituir um auxílio na prevenção de insegurança alimentar, mitigando os riscos sociais e económicos e alcançando a intensificação sustentável nessas condições (FAO, 2017).

Para além do exposto, os meios para atingir a intensificação sustentável necessitam de ser experimentados antes da sua aplicação global, de forma a comprovar a sua mais valia social, económica e ecológica (Garnett et al., 2013).

Em todas as premissas enunciadas a agricultura de precisão afirma-se como uma forma de agricultura que se enquadra no desafio de intensificação sustentável, representando assim uma oportunidade para a realização do objetivo de produzir mais com menos, o que leva a encarar a agricultura de precisão não como uma opção de gestão agrícola, mas sim como uma obrigatoriedade evolutiva do setor primário (Mulla e Schepers, 1997).

Tendo em consideração as diferentes abordagens ao conceito de agricultura de precisão, podemos considerar dois processos distintos quando analisamos esta prática agrícola: a obtenção e análise de informação e, posteriormente, a gestão de informação (Mulla e Schepers, 1997).

A obtenção e análise de informação depende fortemente do nível de investimento tecnológico da exploração, uma vez que para obter uma maior variedade de dados a um nível de especificidade elevado são necessários equipamentos próprios, que geralmente representam um custo importante no momento da sua aquisição. Simultaneamente, é necessária mão-de-obra especializada para a utilização e manutenção dos equipamentos que pode ser adquirida por contratação de prestação de serviços ou por contrato a título singular de um técnico. Qualquer uma das opções referidas acarreta custos importantes (Mulla e Schepers, 1997).

Por outro lado, a gestão de informação pode ser considerada quase totalmente independente do investimento tecnológico de obtenção de dados, pois depende principalmente do conhecimento técnico do agricultor. A interpretação dos dados obtidos e a formação de indicadores é um processo que, embora em muitas ocasiões considerado secundário, é limitante na evolução da agricultura de precisão (Lindblom et al., 2017).

### 1.1.3. Fatores limitantes em agricultura de precisão

Com a acentuada evolução tecnológica nas últimas décadas, juntamente com a redução significativa de preços de sensores e de maquinaria agrícola, anteriormente proibitivos para a realidade da maior parte das explorações agrícolas, estava prevista uma adesão superior à prática de agricultura de precisão. No entanto, a evolução de adesão à agricultura de precisão não correspondeu às expectativas que a tecnologia proporcionou. Uma das causas encontradas para justificar este fenómeno está na incapacidade de o agricultor interpretar e aplicar toda a informação obtida no sistema agrícola, ou seja, uma limitação de digitalização. Esta informação torna-se tão mais abundante quanto mais acentuada é a presença tecnológica capaz de registar e transmitir dados. Os sistemas de suporte à tomada de decisão (DSS, *Decision Support System*) existentes aparentam não lograr a simplificação da decisão do agricultor, mas sim dificultar a mesma atendendo à grande quantidade de informação que se mostram ineficientes em integrar (Lindblom et al., 2017).

No processo gestão de informação encontram-se as fases de relacionamento de dados obtidos, de forma a serem criados indicadores, e também a fase de tomada de decisão. Este processo é o mais limitante aquando da aplicação da agricultura de precisão. Como descrito por Linblom et al. (2017), na Suécia os agricultores tiveram acesso a tecnologias de última geração. No entanto, a transformação em melhorias líquidas na sua produção não era evidente, uma vez que os dados obtidos eram excessivos e os agricultores não integravam o processo de gestão de informação obtida. Primeiramente, foi apontada a falta de conhecimento técnico dos agricultores como a principal causa da inexistência de acréscimos na produção após a aquisição dos equipamentos. Contudo, concluiu-se que, embora o conhecimento dos agricultores fosse uma limitação, a tomada de decisão era baseada em diversos componentes e que os agricultores tinham uma carência, não de informação, mas de como integrar a informação com a sua decisão. Desta forma surgiu a necessidade da presença de DSS's para a integração contínua de uma maior utilização de componentes tecnológicos na tomada de decisão.

Verifica-se que a evolução da adoção da agricultura de precisão está associada à incorporação de equipamentos de elevado nível tecnológico que não necessitam de interpretação analítica para conferirem benefícios nas explorações, como por exemplo, equipamentos de direção assistida ou de condução automática. Esta adoção tecnológica é justificada pela reduzida necessidade de aquisição de novo conhecimento ou de sistemas de auxílio à tomada de decisão para que exista um retorno económico líquido significativo. Pelo contrário, outras tecnologias, como as que dependem de grande quantidade de informação, mesmo não representando investimentos tão avultados, não são facilmente adotadas por não apresentarem benefícios imediatos. Apesar da sua aplicação conduzir a uma melhoria significativa no processo de tomada de decisão, a elevada quantidade de informação e, sobretudo, a complexidade da sua integração constituem um considerável fator limitante para a sua incorporação nas empresas (McBratney et al., 2005).

Um exemplo de tecnologias com a limitação referida é a utilização de imagens culturais que determinem a densidade vegetativa e, por sua vez, possam indicar, após o tratamento dos dados obtidos, o índice de área foliar ou as fases de desenvolvimento em vários pontos da canópia de uma

cultura de milho (Gitelson et al., 2003). Esta tecnologia confere uma grande utilidade na gestão cultural para os produtores de milho. No entanto, o investimento em aquisição de imagens culturais, por si só, não apresenta um benefício líquido na produção, sendo necessário recorrer ao tratamento dos dados obtidos por DSS's, para que o produtor possa retirar algum benefício desse investimento em imagens culturais.

O comportamento evolutivo da adoção de agricultura de precisão, por parte dos agricultores, parece estar bastante restringido a tecnologias que, quando adquiridas, apresentem um retorno imediato, em detrimento de tecnologias menos dispendiosas, mas que exijam um tratamento dos resultados obtidos para a obtenção de retorno. Torna-se assim evidente que a informação disponível pode ser um fator limitante quando em grandes volumes, variedade ou velocidade de aquisição (McBratney et al., 2005).

#### **1.1.4. *Big Data***

“*Big Data* é a designação utilizada quando o volume de dados, a velocidade de aquisição, ou a representatividade dos dados [variedade] limitam a habilidade de executar uma análise eficiente através de abordagens tradicionais ou que para tal seja necessário recorrer a redimensionamento dos dados” (Cooper e Mell, 2012).

O termo *Big Data* teve origem em meados da década de 1990, respondendo à necessidade de designação da gestão e processamento de bases de dados de grandes dimensões. No entanto, *Big Data* engloba outros atributos para além das dimensões das bases de dados, ou seja, não basta que uma base de dados apresente milhares de registos para ser categorizada de *Big Data*, mas sim um conjunto de características que conferem uma maior complexidade à sua análise quando em comparação com bases de dados convencionais (Laney, 2001).

Numa abordagem inicial para definir *Big Data*, Laney (2001) distingue bases de dados convencionais de *Big Data* pela regra dos 3 V's. Esta regra determina que uma base de dados corresponde a *Big Data* se os seus dados possuírem Velocidade, Volume e Variedade. Por outras palavras, segundo esta categorização, *Big Data* corresponde a uma base de dados em que os dados sejam atualizados regularmente e o mais próximo possível da atualização em tempo real, mas também corresponda a um elevado volume de registos e a uma variedade estruturada ou não nos dados presentes na base de dados.

No entanto, com o desenvolvimento tecnológico e informático recente, a definição de *Big Data* foi revista, conferindo mais atributos para a classificação de bases de dados em *Big Data*, assim como imputando ordem de importância aos atributos que definem as bases de dados em função do quão limitantes são para a análise das mesmas. Num contexto mais atual, *Big Data* é definido pelos 3 V's mas também pela exaustividade, resolução, relacionamento e escalabilidade dos dados (Kitchin, 2016).

Em concreto, a exaustividade de uma base de dados diz respeito à capacidade da base de dados representar a totalidade do sistema; a resolução implica baixos níveis de granularidade dos



dados existentes; o relacionamento implica a existência de atributos comuns em diferentes conjuntos de dados que permitam o seu relacionamento numa base de dados e a escalabilidade diz respeito à capacidade da base de dados em aumentar as suas dimensões celeremente consoante as necessidades (Kitchin, 2016).

Finalmente, os atributos enunciados para a distinção entre bases de dados convencionais e *Big Data* possuem importâncias distintas na classificação, sendo que, dos atributos da Tabela 1, a velocidade e a exaustividade constituem os mais importantes, enquanto o volume e a variedade representam os menos relevantes para a definição como *Big Data*. Esta categorização dos atributos é justificada pela dificuldade acrescentada à análise da base de dados que cada um apresenta. A velocidade de atualização de dados e a sua exaustividade representam maiores fatores limitantes para a análise da base de dados quando em comparação com o volume da base de dados. Uma vez que o poder de processamento informático existente é cada vez maior, o volume da base de dados toma uma posição menos limitante no momento da sua análise. Em contraste, a velocidade de atualização e a exaustividade dos dados representam um desafio no constante e abrangente processamento dos dados (Kitchin, 2016).

Tabela 1 - Comparação entre bases de dados convencionais e bases de dados de Big Data, segundo Kitchin (2016).

	<b>Convencional</b>	<b><i>Big Data</i></b>
<b>Volume</b>	Limitado a elevado	Muito elevado
<b>Velocidade</b>	Lento a parado	Rápido e contínuo
<b>Variedade</b>	Limitado a abrangente	Abrangente
<b>Exaustividade</b>	Amostras	Populações inteiras
<b>Resolução</b>	Grosseira a fina	Fina
<b>Relacionamento</b>	Fraca a forte	Forte
<b>Escalabilidade</b>	Pequena a média	Elevada

#### 1.1.5. *Big Data* agrícola

Como anteriormente referido, o desenvolvimento e utilização de tecnologias agrícolas em agricultura de precisão e agricultura convencional estão diretamente relacionado com o aparecimento de um maior volume e variedade de dados, a uma velocidade cada vez mais próxima da atualização em tempo real. Por outras palavras, este desenvolvimento está a gerar um aparecimento de *Big Data* agrícola (Sonka, 2016).

No setor agrícola, o *Big Data* é encontrado em diferentes formas tecnológicas, seja, por exemplo, na forma de sensores ambientais utilizados em equipamento agrícola e que cedem os seus dados de forma relativamente simples para um local de armazenamento, seja na forma de algoritmos de aprendizagem profunda, como a construção de uma estratégia de proteção cultural, ou ainda em fontes exteriores à exploração, como indicadores provenientes de estações climáticas, indicadores de preços de mercado ou mesmo indicadores provenientes de outras explorações (Wolfert, 2017).

O *Big Data* agrícola pode ser obtido de forma estruturada ou destruturada, sendo possível distinguir a tipologia de dados obtidos em seis diferentes categorias (Bronson et al., 2016):

- Dados históricos, onde se incluem análises de solos, monitorização de colheita, condições climáticas, dados geográficos e mão-de-obra;
- Dados sensoriais e equipamento agrícola. Neste grupo, encontram-se dados recolhidos por equipamentos de deteção remota, equipamentos GPS, registos de equipamentos, fertilizadores VRT (*Variable Rate Fertilizers*), ou sensores de humidade e de temperatura;
- Dados sociais e dados originários da internet, como dados de satisfação de cliente, sítios na Internet com informação agrícola, redes sociais e dados de motores de pesquisa;
- Publicações. Neste grupo, incluem-se parâmetros provenientes de investigação científica agrónoma e de bibliografia que inclua indicações e requerimentos culturais;
- Dados provenientes de *streams* de informação. Aqui estão incluídos grupos com origem em fontes contínuas de dados como monitorização cultural através de drones, aviões, *smartphones*, sistemas de segurança, entre outros;
- Dados externos. Finalmente, neste grupo estão incluídos dados provenientes de terceiros como informações governamentais, outras empresas agrícolas ou a indústria de produção de equipamentos agrícolas.

#### **1.1.6. Agricultura de precisão como precursora para a utilização de *Big Data* agrícola**

Como referido anteriormente, a agricultura de precisão possui um forte enfoque no processamento de dados, sendo que esta área da agricultura de precisão chega mesmo a ser um fator limitante para os benefícios da sua utilização numa exploração agrícola. Esta limitação é criada sobretudo pelo volume, variedade e velocidade de aquisição de dados, à medida que são introduzidos na exploração novos equipamentos fornecedores de dados (Gitelson et al., 2003; Lindblom et al., 2017).

Embora os conceitos de *Big Data* e de agricultura de precisão não sejam sinónimos, a natureza da definição de agricultura de precisão confere um substancial acréscimo de dados disponíveis para o agricultor respeitantes à sua produção. Pelo processo produtivo agrícola ser dinâmico e sujeito a variáveis voláteis – como a precipitação, radiação ou temperatura – é inevitável que a tendência das bases de dados de agricultura de precisão tendam para o formato de *Big Data* (Sonka, 2016).

Os dados originários da agricultura de precisão atual respeitam a definição de *Big Data* dos 3 V's, uma vez que, ao ser necessária uma intervenção diária e célere na vasta maioria das operações agrícolas, os dados são atualizados com a maior periodicidade possível e idealmente em tempo real, conferindo assim a velocidade de aquisição de dados que define *Big Data*. Por outro lado, tratando-se de um sistema complexo, para incorporar o sistema de produção agrícola numa base de dados são necessários diferentes tipos de dados culturais, edafoclimáticos e socioeconómicos, conferindo assim uma elevada variedade aos dados produzidos. Por fim, sendo um sistema biológico, encontra-se em constante mudança no que diz respeito aos registos realizados, pelo que também é necessário um substancial volume de dados para uma tomada de decisão representativa das reais necessidades do

sistema. Em suma, embora agricultura de precisão e *Big Data* sejam conceitos distintos, encontram-se indubitavelmente relacionados, sendo que para um aproveitamento eficiente das práticas de agricultura de precisão é necessário assumir os dados produzidos como *Big Data* e processar os mesmos como tal (Sonka, 2016).

#### **1.1.7. Importância e oportunidades agrícolas na utilização de *Big Data***

A utilização de *Big Data* por parte das organizações agrícolas permite o alargamento dos objetivos e focos por parte das explorações agrícolas, uma vez que ao abranger mais informação sobre o sistema onde está inserido permite um aumento de eficiência, não só na vertente produtiva, mas também em vertentes ecológicas e ambientais já referidas anteriormente, tais como a segurança alimentar e a sustentabilidade. Por outro lado, o foco da utilização de *Big Data* no setor primário não abrange somente a produção agrícola, mas sim toda a cadeia de valor, desde o setor primário até ao consumidor, quer pela produção mais eficiente, quer pela possibilidade integração das externalidades no processo produtivo das organizações agrícolas (Wolfert, 2017).

Uma importante oportunidade para um maior aproveitamento do *Big Data* agrícola e consequente digitalização da agricultura é a IoT (*Internet of Things*). A IoT permite a conexão de vários equipamentos tecnológicos entre eles e à base de dados da exploração, assim como a conexão a fontes de dados exteriores à exploração. Este fenómeno significa que as organizações agrícolas conseguirão gerar e integrar dados relevantes e em tempo real no seu processo produtivo, o que eventualmente se traduzirá numa maior eficiência na utilização dos seus recursos – sejam eles físicos ou financeiros – através do tratamento de *Big Data*, avaliando a situação atual, os seus registos históricos e mesmo realizando previsões para tomadas de decisões o mais eficiente e eficazes possíveis, com os dados disponíveis (Wolfert, 2017). Por outro lado, com um processamento adequado do *Big Data* adquirido, os próprios equipamentos poderão ajustar as suas funções de acordo com as informações provenientes de outros, em tempo real, uma vez mais aumentando a eficiência dos processos produtivos agrícolas (Sonka, 2016).

Embora atualmente seja simultaneamente um desafio e uma limitação, a gestão analítica dos dados é também uma das mais relevantes oportunidades para a utilização de *Big Data* agrícola. Dados relativos a fontes de informação convencionalmente não utilizados como dados de redes sociais, dados provenientes de satélites e máquinas automatizadas na forma de imagens ou vídeos, representam uma fonte de *Big Data* importante que poderá vir a ser explorada por organizações agrícolas. No caso das redes sociais, as informações disponíveis podem, por exemplo, transmitir informações relevantes sobre o consumidor final e permitir a alteração de características dos produtos produzidos em função dessa informação. Para isso é necessário a capacidade analítica evoluir, uma vez que os dados já se encontram disponíveis na atualidade. No caso de dados em forma de vídeo ou imagem, a tecnologia de produção de dados já existe, e a limitação é, uma vez mais, a capacidade de tratamento analítico. Em suma, a barreira de processamento analítico representa uma limitação, mas também uma forte oportunidade de progresso para uma agricultura que utilize o *Big Data* disponível no seu sistema (Wolfert, 2017).

### 1.1.8. Desafios do *Big Data* no setor primário

O surgimento de possibilidades de utilização e processamento de *Big Data* no contexto de produção agrícola levanta algumas limitações presentes e futuras que representarão desafios a ultrapassar (Wolfert, 2017).

Limitações relacionadas com a propriedade dos dados produzidos, e a sua privacidade e segurança. Em que embora já se encontrem a ser legislados e regulados alguns aspetos no contexto nacional e internacional, ainda é necessário maior salvaguarda dos interesses dos proprietários dessa informação, sem com isso abrandar os desenvolvimentos tecnológicos que permitem o avanço da valorização da informação (Wolfert, 2017).

Um outro desafio da utilização de *Big Data* abrange a qualidade dos dados obtidos e a sua representatividade, uma vez que ao ser tratado um elevado volume de registos (a uma elevada velocidade e com variedade nos dados obtidos) a tendência é o aparecimento de dados incorretos dentro das bases de dados, o que obriga à criação de mecanismos de controle de dados adaptados aos contextos específicos de cada organização, não existindo para já uma solução universal para todas as organizações, uma vez que os contextos produtivos se alteram em cada situação empresarial particular (Wolfert, 2017).

Em seguimento do desafio anterior, o processamento analítico eficiente e adequado a cada organização requer a junção do conhecimento de ciência analítica, ao conhecimento da área em que a organização opera. Isto é um fator bastante limitante para qualquer base de dados eficiente no contexto organizacional onde se encontra. Esta limitação é relevante e representa um desafio, uma vez que para uma base de dados ser otimamente construída e posteriormente processada é necessária a junção de duas áreas: uma transversal às organizações e outra específica à atividade da organização. Somente desta forma se conseguirá um processamento de *Big Data* eficiente e adequada às necessidades de cada organização em particular. Para ultrapassar este desafio será necessário uma relação inteligente entre um especialista em ciência analítica e um especialista da área operacional da organização (Wolfert, 2017).

A integração de *Big Data* no modelo de negócios de uma empresa necessita de ser realizada de uma forma sustentável, o que representa um outro desafio, uma vez que, por um lado, o modelo de negócios da organização beneficia com o máximo de informação relevante processada para os seus processos produtivos mas, por outro lado, uma transição demasiado célere para uma dependência do processamento de *Big Data* causará um risco em que alguns aspetos, anteriormente não contemplados, possam vir a revelar fragilidades na forma de processamento desta informação. Torna-se necessário um estudo extensivo das particularidades dos processos organizacionais para a introdução integrada da utilização de *Big Data* nesses mesmos processos (Wolfert, 2017).

Finalmente, outro desafio na utilização de *Big Data* será o desenvolvimento de plataformas abertas para o desenvolvimento de soluções e inovações no que diz respeito a diferentes formas de processamento de dados com as características da *Big Data* – elevados volume, velocidade e

variedade (Laney, 2001) – de forma a acelerar inovações, mas também para conferir ao setor primário uma posição mais consolidada na cadeia de valor (Wolfert, 2017).

## **1.2. Sistemas de apoio à tomada de decisão (DSS)**

### **1.2.1. Tomada de decisão**

A tomada de decisão está presente no nosso quotidiano enquanto indivíduos, quer no âmbito pessoal, quer no âmbito profissional, assim como em entidades coletivas, como empresas ou associações. A ação de tomada de decisão consiste em conhecer o problema, avaliar a importância de uma decisão, incorporar os critérios em que a tomada de decisão assenta, as suas consequências e comparar ações alternativas (Shim et al., 2002).

Numa primeira abordagem crítica à qualidade da decisão a tomar, a quantidade de informação disponível parece ser diretamente proporcional à qualidade da decisão final. Tal não é o caso, uma vez que as informações em que os passos para a tomada de decisão se alicerçam não são hierarquicamente iguais em termos de importância para a resolução de um problema. Este fenómeno traduz-se na necessidade de interpretação de uma tomada de decisão como um processo que engloba vários critérios (Saaty, 2008).

Quanto mais complexo é o problema original em que a tomada de decisão influi, maior o número de variáveis que o processo de tomada de decisão necessita considerar. Tendo em conta a necessidade de incorporar vários critérios e informações, a análise de decisão multicritérios (MCDA, *Multi-Criteria Decision Analysis*) torna-se fundamental para a obtenção de uma decisão integrada no sistema onde o problema se encontra. A MCDA é uma área que estuda a tomada de decisão enquadrada nas variáveis presentes no sistema em análise.

Existem diversas formas de abordar um processo de tomada de decisão, por outras palavras, existem vários processos de MCDA, todos com as suas valências e limitações, amplamente estudados, que abordam os problemas sempre numa ótica de integração de critérios e que têm como característica usualmente poderem ser incorporados em sistemas informáticos de processamento. A junção de processos MCDA e sistemas informáticos de processamento origina os DSS's (Wang et al., 2009).

### **1.2.2. Processo de tomada de decisão**

O processo de tomada de decisão quando em estudo por uma MCDA, contempla cinco fases para obter a melhor solução para o problema em questão: definição do problema de decisão, identificação das alternativas potenciais, seleção de fatores e critérios, comparação das diferentes alternativas e, finalmente, escolha da melhor alternativa (FAO, 2017).

Na primeira fase do MCDA é pretendida a obtenção clara dos objetivos da tomada de decisão a realizar. Por exemplo, num problema de calendarização de colheita num pomar com diferentes variedades e diferentes épocas de maturação do fruto, o objetivo da tomada de decisão será a organização das diferentes variedades em função da sua época de maturação.

Durante a fase de identificação de alternativas potenciais, são recolhidas as opções ou escolhas que se podem adequar à situação a abordar.

A fase de seleção de fatores e critérios de decisão consiste no estabelecimento de hierarquias entre os critérios que afetam a decisão. Num exemplo agrícola, podemos considerar, uma vez mais, um pomar e a tomada de decisão da data de colheita, que neste exemplo dependerá somente do teor de sólidos solúveis no fruto e de uma curta janela de mercado. Neste exemplo, o critério da curta janela de mercado será mais importante hierarquicamente do que o teor de sólidos solúveis, uma vez que se não se vender o produto durante o curto período, não será relevante que o teor de sólidos solúveis esteja nos parâmetros ótimos pois não existe comprador para o produto final.

A comparação de alternativas consiste na análise de dados das várias alternativas encontradas para a tomada de decisão. Nesta fase é estudado o resultado esperado de cada alternativa e posteriormente cruzados os resultados de cada uma.

Por fim, é realizada a escolha da melhor alternativa para a tomada de decisão, baseada na análise realizada anteriormente (FAO, 2017; Shim et al., 2002).

### **1.2.3. Programas MCDM**

Com a crescente oferta tecnológica, os programas de MCDM (*Multi-criteria Decision Making*), mais concretamente os sistemas de informação geográfica MCDM, evoluíram bastante ao longo do século XX e continuam o seu desenvolvimento na atualidade, representando um acentuado avanço nas potencialidades da agricultura de precisão. Esta tecnologia permitiu adicionar várias camadas de informação a uma única parcela referenciada geograficamente, resultando num leque mais amplo de informações disponíveis para o agricultor, podendo este relacionar dados de várias ordens, como características edafoclimáticas, características de rega e fertilização, e características culturais, como anteriormente não era possível. Consequentemente, o processo de tomada de decisão torna-se bastante mais completo e englobante, porém, mais complexo. (Chen, Yun et al., 2010).

Efetivamente, a tomada de decisão no contexto agrícola deriva da inclusão de variados fatores no processo. Alguns fatores, como as características edafoclimáticas locais, características socioeconómicas em que a exploração se insere, características culturais das espécies utilizadas, problemas de proteção cultural, representam variáveis que os agricultores têm de ter em consideração para a obtenção de uma decisão final integrada no seu problema. Todos estes fatores são comuns aos agricultores, embora o seu conteúdo seja variável para cada situação em concreto, ou seja, não existe um manual universal que logre apresentar soluções à escala da exploração, para todos os agricultores,

partindo assim destes a contextualização do seu problema para tomar uma decisão em função dos seus critérios e variáveis locais (Stafford, 2000).

Face à complexidade da atividade agrícola com variáveis de difícil interpretação e previsão, muitas das decisões tomadas pelos agricultores são baseadas em intuição, uma vez que existe uma janela de oportunidade para estas serem tomadas, não permitindo ao agricultor um estudo aprofundado da situação em que se encontra para obter todas as informações necessárias para uma tomada de decisão completamente fundamentada. A intuição dos agricultores consiste num conciliar do conhecimento empírico e académico adquirido, juntamente com a interpretação da situação em que se encontram e da, normalmente, curta janela de oportunidade de intervenção (Suter, 1992). Este tipo de decisões intuitivas são uteis em situações em que o agricultor necessita de intervir rapidamente e a não intervenção causa necessariamente uma pior situação.

O fator tempo desempenha, por conseguinte, um papel importante no tipo de tomada de decisão. Devido à complexidade de variáveis presentes no setor agrícola, nem sempre é possível para um agricultor estruturar o seu processo de decisão contemplando todas as informações disponíveis. Apesar de existirem formulações de modelos que otimizam a interpretação de dados, a limitação para estes modelos reside no tempo disponível e na velocidade de processamento de que o agricultor dispõe. Em suma, são conhecidas formas de otimização do processo de tomada de decisão agrícola, mas que, no entanto, não são utilizadas em virtude do tempo requerido para a execução das mesmas que contrasta com a necessidade de ação célere por parte dos agricultores (Fountas et al., 2006).

Surge assim a necessidade de incorporar o conhecimento agrícola do agricultor a um sistema que consiga processar atempadamente o volume de dados disponíveis a uma escala de nível específico, de forma a permitir uma visão holística do sistema em que a exploração agrícola está inserida. Por outras palavras, surge a necessidade de um sistema de apoio à tomada de decisão, DSS (Decision Support System).

“Os sistemas de apoio à tomada de decisão (DSS) são soluções de tecnologia informática que podem ser utilizadas para suportar tomadas de decisão e resolução de problemas” (Shim et al., 2002).

Com o constante desenvolvimento tecnológico na atualidade, inclusivamente em tecnologias e técnicas agrícolas, verifica-se que a principal limitação ao desenvolvimento da agricultura como um todo, não é a falta ou dificuldade ao acesso de dados agrícolas e económicos em tempo real numa escala geográfica específica e reduzida, mas sim a forma de interpretar e processar os dados obtidos de forma a otimizar a eficiência agrícola no uso de informação e, consequentemente, otimizar a eficiência do uso de insumos e melhorar os benefícios líquidos (Stafford, 2000).

Tendo em consideração as limitações apontadas, a agricultura de precisão enfatiza cada vez mais a gestão sobre a análise dos dados obtidos, em detrimento da separação entre análise de dados e gestão agrícola. A integração da análise de dados na gestão de uma exploração agrícola representa uma problemática importante, uma vez que um agricultor não consegue realizar sem ferramentas de processamento. Deste impedimento nascem as ferramentas de apoio à tomada de decisão (DSS – Decision Support Systems) que funcionam como programas de receção e processamento de

informação, de forma a que o utilizador final consiga tomar decisões baseadas no máximo de informações que dispõe (Mamo et al., 2003).

Segundo Adamchuk et al. (2004), para melhorar o conhecimento de atributos agronómicos do solo à escala da parcela e a uma escala específica mais reduzida é necessária a integração de várias fontes de dados disponíveis. Esta fusão de dados só é possível, a escalas de localização específica, através de sistemas de processamento de dados, pelo que com o aparecimento de novos sistemas de apoio à decisão, este conhecimento se torna acessível.

Por outro lado, os sistemas de apoio à decisão são também relevantes no estudo de registos históricos de grandes volumes de dados. Estes registos históricos podem servir de meios para atingir melhores opções quanto ao controlo cultural e a otimização de estratégias culturais, assim como no estudo da variabilidade temporal de fatores ambientais, mais uma vez, a uma escala espacial reduzida (Hokazono e Hayashi, 2012).

#### **1.2.4. Utilidade e funcionamento dos DSS**

À semelhança do processo de análise multicritérios da tomada de decisão, Shim et al. (2002) descrevem o processo de funcionamento de um sistema de apoio à decisão em sete fases principais, designadamente, identificação do problema, definição do problema, geração de alternativas, desenvolvimento do modelo, análise alternativa, escolha, implementação.

Posteriormente às fases de identificação e definição do problema, algumas alternativas são elaboradas, dando origem a diferentes modelos para a análise das implicações de implementação das diferentes soluções. Por fim, existe um processo de decisão entre as alternativas que, por sua vez, define a implementação da solução encontrada (Shim et al., 2002).

Para uma ferramenta clássica de apoio à tomada de decisão ser funcional, a mesma deverá apresentar três componentes principais: uma base de dados, capacidade de processamento e uma interface de utilizador (Shim et al., 2002).

Para um apoio à decisão representativo, a ferramenta terá de dispor de uma base de dados representativa, bem como de um grau de especificidade suficiente para responder ao grau de especificidade da solução pretendida. Esta base de dados deverá ter acesso a dados internos e externos correspondentes ao problema a abordar, à informação sobre o problema e às capacidades de resolução do mesmo, e acesso a conhecimento de forma a tratar os dados e as informações disponibilizadas (Shim et al., 2002).

A capacidade de processamento é um dos aspetos principais de um DSS, uma vez que o objetivo do mesmo é o relacionamento de um grande volume de dados, de uma forma que não seria possível ser realizada por um indivíduo, num espaço de tempo útil (Shim et al., 2002).

Finalmente, os resultados de um DSS devem ser apresentados numa interface de utilizador inteligível, o que significa que deve lograr a interpretação do seu utilizador, sem que com isso perca a



representatividade dos seus resultados, através de gráficos, tabelas e interatividade entre o DSS e o utilizador (Shim et al., 2002).

### 1.2.5. Diretrizes futuras dos DSS's

Segundo McBratney et al. (2005), o futuro da agricultura de precisão passa pelo desenvolvimento de ferramentas de apoio à tomada de decisão. Essas ferramentas devem contemplar funções como a utilização e integração de novas tecnologias instaladas na exploração, realizar análises de custo/benefício comercial em áreas específicas da parcela onde está instalada a cultura, integrar tecnologias de forma a obter desde uma visão holística da exploração, a uma especificidade local cultural e, finalmente, estabelece protocolos distintos para uma abordagem de gestão localmente específica em diferentes culturas. Em suma, as ferramentas de apoio à tomada de decisão não deverão ser somente ferramentas de indicadores culturais, mas sim englobar as características dos processos a realizar e incorporar a dimensão económica na altura de prestar o apoio à tomada de decisão.

Numa maior escala temporal, Verdouw et al. (2015) detalham a composição de funções de um sistema automatizado de tomada e execução de decisão, sendo que as funções descritas se encontram por ordem de implementação, ou seja, a última função deste sistema será também a mais distante temporalmente da sua implementação prática. Esta ideia associa a agricultura de precisão e a *smart farming* à utilização de *Big Data* (posteriormente abordado). As funções deste sistema são (Figura 1):

- Utilização de sensores e monitorização dos mesmos. Nesta função é pretendido avaliar o desempenho de processos ao nível da exploração, o que pode ser realizado manualmente ou automaticamente com recurso à utilização de tecnologia como sensores ou satélites, assim como utilização de dados externos para complementar os dados obtidos;
- Análise e tomada de decisão. Esta função seria idêntica ao processo de tomada de decisão descrito por Shim et al. (2002), em que também se incluiria a comparação dos dados obtidos com normas que especifiquem o resultado esperado e a correção de erros ou dados não representativos, de forma a tomar uma decisão o mais eficiente possível de acordo com os dados disponibilizados pela função anterior;
- Intervenção. Finalmente, a função mais complexa a atingir seria a execução, que consistiria em planear e implementar a decisão automaticamente na exploração, de forma a corrigir ou mesmo realizar totalmente processos da exploração, como por exemplo, a atribuição de novas ordens de trabalho a funcionários da empresa (Wolfert, 2014).

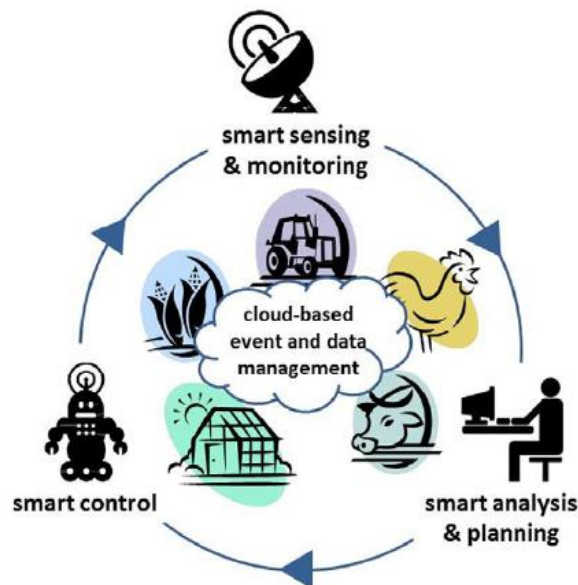


Figura 1 – Ciclo de gestão de uma organização agrícola em regime *Smart Farming* alicerçada numa base de dados *online* e num DSS (Wolfert et al., 2014).

Nesta dissertação a abordagem à agricultura de precisão incide sobre o que McBratney et al. (2005) identificaram nos sistemas de apoio à tomada de decisão como fatores limitantes e a melhorar no futuro, representados pela contemplação de processos e pela dimensão económica como partes fundamentais de um sistema de apoio à decisão.

#### 1.2.6. Internet como ferramenta

Com a afirmação da internet como plataforma comum de diferentes tecnologias e utilizadores, é hoje possível aceder a ferramentas de apoio à tomada de decisão numa grande variedade de dispositivos e em qualquer zona que possua cobertura de rede móvel de telecomunicações, o que possibilita a criação de DSS's para uma maior diversidade de utilizadores finais. Por essa razão, a evolução do desenvolvimento de DSS's contempla uma maior dinâmica do sistema, permitindo uma maior diversidade de utilizadores finais com diferentes funções na exploração, e uma menor formação para a utilização do DSS. A internet abre também uma janela de oportunidade de colaboração entre diferentes membros da empresa e entre vários equipamentos através do mesmo DSS.

A crescente disponibilidade de cobertura “wifi” e 4G-LTE permite não só o acesso a dados e a indicadores por parte de operadores, agricultores e de máquinas, mas também a obtenção de dados de várias fontes, quer fontes dentro da exploração, como sensores ou inputs de operadores, quer fontes externas, como imagens de satélites ou previsões meteorológicas.

O acesso à internet, por parte de equipamentos e objetos, e a sua capacidade de partilhar informação entre várias plataformas, num formato compatível, é denominado por IoT (Gubbi et al., 2013). Esta tecnologia, quando aplicada à agricultura e mais concretamente aos sistemas de auxílio à

tomada de decisão, abre novas fronteiras no que diz respeito à forma e ao tipo de dados que podemos obter e fornecer ao DSS.

Ao ser possível integrar informação de várias fontes na exploração num único sistema, em tempo real, a quantidade de dados disponíveis para utilização no processo de tomada de decisão aumenta significativamente. Com esse aumento e em conjugação com ferramentas de processamento, a precisão e eficiência da utilização de insumos aumenta também. Existe assim a possibilidade de criação de um novo modelo de gestão de agricultura de precisão que passa pela utilização de toda a informação disponível, de forma a obter, não só registos históricos, mas também informação georreferenciada em tempo real. Esta informação pode consistir somente em indicadores técnicos, mas pode incluir também informações financeiras, económicas e operacionais igualmente em tempo real (Ye et al., 2013).

Ao desenvolvimento destas novas formas de utilização de DSS's, que conferem um maior nível de especialização a cada situação particular em que a exploração agrícola se encontra, está associado um benefício líquido, uma vez que a criação de interfaces de utilizador personalizadas pela função do utilizador final e em função de cada parcela e cultura que a exploração detém, permite a criação de valor em forma de informação relevante para os processos produtivos, com um investimento reduzido (Shim et al., 2002).

Como referido, a internet possibilita um maior número de insumos de informação que são anexados à base de dados de um sistema. Um exemplo dentro da exploração da potencialidade da ligação permanente à internet é dado pelo WSN (*Wireless Service Network*). O WSN consiste numa rede de sensores, por definição não dispendiosos, com a capacidade de registar e transmitir dados edafoclimáticos para uma estação que, posteriormente ao recebimento dos dados, transmite a informação para um sistema de processamento de dados. Esta rede de sensores representa uma importante oportunidade para a análise de dados, uma vez que, a partir destes, pode ser alcançada uma vasta quantidade de dados relevantes para a quantificação de indicadores ambientais, como por exemplo, a precipitação, a temperatura, a incidência de radiação, entre outros, bem como a sua variabilidade espacial ao longo da parcela em estudo (Vieira et al., 2003).

Com o desenvolvimento da capacidade de processamento de dados, as WSN são um dos pontos onde os DSS estão a evoluir na atualidade. Num passado recente, o volume de dados produzidos seria inconcebível de processar. No entanto, a evolução aponta para sistemas em que é formada uma rede de indicadores ambientais e culturais georreferenciados que servem de base para várias camadas de informação. Por outras palavras, o utilizador final desta informação, consegue obter, em tempo real, informações relativas aos indicadores ambientais e culturais e tomar decisões, aplicando as mesmas remotamente, com uma elevada especificidade, como por exemplo alterar a dotação de rega num determinado setor ou alterar a quantidade de fertilizante a aplicar num local específico, aumentando assim a produtividade e a qualidade final dos seus produtos, e simultaneamente aumentando a eficiência de utilização de insumos (Srbinovska et al., 2015).

### 1.2.7. Programas *Business Intelligence* no tratamento de *Big Data*

Um sistema de *Business Intelligence* (BI) consiste num “conjunto de ferramentas, tecnologias e produtos de programação que são utilizados para adquirir, integrar, analisar e disponibilizar dados” (Reinschmidt, 2000). Por outras palavras, a principal tarefa de um sistema de BI inclui “a exploração inteligente, integração, agregação e uma análise multidimensional proveniente de várias fontes de dados” (Olszak, 2007).

Nesta definição de sistemas de BI encontra-se implícito que os dados são tratados como um recurso altamente valioso para as empresas e que a transformação de uma quantidade de dados para a qualidade dos mesmos é crucial. Como tal, elevados volumes de dados de uma grande empresa podem ser transformados numa forma coerente para conferir uma visão holística da situação em que a empresa se encontra. Por esta razão, os sistemas BI são também definidos por permitirem a entrega de informação relevante na altura, no local, e na forma correta, de maneira a que consigam auxiliar indivíduos, departamentos ou mesmos a empresa como um todo, no momento de tomada de decisão (Yeoh, 2010).

Os sistemas de BI são usualmente implementados para substituir programas de planeamento de recursos de empresas, designados por ERP (*Enterprise Resource Planning*), uma vez que permitem a execução de tarefas anteriormente realizadas pelos ERP e acrescentam novas capacidades à empresa relacionadas com a análise célere e específica de grandes volumes de dados, incluindo o *Big Data* abordado anteriormente. A escala de investimento e de implementação destes sistemas tem vindo a aumentar globalmente, refletindo a sua importância nos processos das empresas e aumentando a necessidade de casos de estudo nas várias áreas em que pode ser implementado (Elbashir, 2007).

No período inicial do seu aparecimento, os sistemas de BI eram vistos apenas como ferramentas de apoio estratégico à tomada de decisão em grandes empresas de serviços. No entanto, com o desenvolvimento recente destes sistemas, o aproveitamento das oportunidades que estes conferem, alargou o seu espectro de ação e, recentemente, os sistemas de BI começam a ser utilizados numa área mais vasta de setores de produção (Rogge, 2005). As empresas utilizam agora estes programas na otimização de processos operacionais, no melhoramento da eficiência da cadeia de valor, na produção e no apoio ao cliente. Este fenómeno deve-se à possibilidade de os sistemas de BI conferirem uma grande variedade de benefícios operacionais ao longo da cadeia de valor (Elbashir, 2008).



Numa vertente tecnológica, a Herdade Vale da Rosa detém e produz registos informáticos de todas as suas operações e consumos, assim como regista indicadores climáticos e de qualidade do seu produto de forma a obter informação para uma otimização da sua produção. Dados relativos à constituição geográfica da sua exploração também se encontram no leque de dados produzidos pela empresa, sendo que todos esses dados se encontram categorizados quanto às suas dimensões geográficas, mas também quanto às suas características culturais, como por exemplo a variedade existente em cada quadra.

Com o forte investimento na componente tecnológica da produção de uva de mesa, a Herdade Vale da Rosa encontra-se com um problema, já descrito como limitante (Lindblom et al., 2017), que está relacionado com a incorporação de equipamentos que acrescentam valor aos processos produtivos na forma de dados. Neste caso em concreto, a empresa possui e produz diariamente uma significativa quantidade de dados provenientes dos diversos equipamentos que utiliza nos seus processos, possibilitando assim afirmar que a Herdade Vale da Rosa é produtora de *Big Data* agrícola, uma vez que a sua produção de dados possui um elevado volume, variedade e velocidade (Boyd e Crawford, 2012; Kitchin, 2013).

A produção de *Big Data* por parte da Herdade Vale da Rosa, por si só, não representa um problema. O verdadeiro factor limitante reside na interpretação desses mesmos dados. A não utilização de toda a informação produzida, impossibilitará a exploração de ser o mais eficiente possível no seu processo produtivo.

Neste caso de estudo, a solução proposta para ultrapassar a limitação enunciada foi a utilização da metodologia de um serviço da CONSULAI, o AGRO BI®. Este serviço consiste na integração de dados de múltiplas fontes e conhecimento técnico agrícola, com situações empresariais particulares, permitindo a análise de dados para cada situação empresarial específica.

O AGRO BI® faculta um serviço de apoio empresarial que permite a gestão e processamento de grandes quantidades de informação, de forma a que o resultado final seja acessível e relevante para toda a empresa. Este serviço utiliza programas de tratamento de dados como o Power BI® e o Microsoft Excel®, permitindo a combinação e análise da informação num só ambiente de trabalho, por um ou vários utilizadores. Estes programas permitem que o serviço AGRO BI® consiga transformar e relacionar diferentes fontes de dados, possibilitando uma tomada de decisão mais eficiente e informada.

Entre as diversas valências do serviço AGRO BI®, são de destacar a avaliação e diagnóstico da exploração, a interpretação e integração dos processos particulares de cada empresa, a criação de indicadores relevantes e o acompanhamento da evolução da empresa sempre com o enfoque de que o serviço se encontre ajustado às necessidades reais da mesma.

## **2.2. Objetivo geral**

Este caso de estudo tem como objetivo principal a criação e validação prática de modelos de *Business Intelligence* capazes de responder aos principais desafios da análise de *Big Data* no processo de tomada de decisão na empresa Herdade Vale da Rosa.

## **2.3. Objetivos específicos**

### **2.3.1. Análise de dados**

Uma vez que a exploração detém uma vasta diversidade de dados relacionados com consumos, operações, condições edafoclimáticas, indicadores de qualidade do produto e também dados financeiros, um dos objetivos deste trabalho consiste na análise dos dados a utilizar em função dos resultados a obter com cada modelo. Estes deverão ser englobantes, mas sem perder o seu foco, que será sempre a resolução de uma limitação associada à interpretação de um ou mais grupos de dados.

### **2.3.2. Implementação de modelos**

A implementação do modelo corresponde a um outro objetivo importante neste caso de estudo. Este objetivo implica não só a incorporação do objetivo anterior, mas também a formulação de modelos que consigam ser sustentáveis e fidedignos à medida que novos dados vão sendo produzidos durante diferentes campanhas agrícolas. Isto significa que os modelos necessitam de possuir a capacidade de continuarem relevantes na resolução de limitações associadas à interpretação de dados numa vasta gama de variações que os mesmos possam vir a sofrer.

Este objetivo ultrapassa a criação de modelos, uma vez que se pretende que os mesmos sejam utilizados por profissionais da empresa, o que implica a facilidade de incorporação de novos dados no modelo e também a facilidade de utilização e interpretação do mesmo.

### **2.3.3. Verificação**

O objetivo de verificação consistirá na determinação de validade da utilização dos modelos criados baseados em software de BI nos vários contextos onde foram aplicados, nomeadamente no tratamento de dados de operações e de consumos de insumos.

Neste objetivo de verificação é também pretendido o estudo da relevância da integração de modelos de BI, na vertente de processos produtivos assim como na vertente técnica agronómica, registando as limitações encontradas, assim como os benefícios da utilização destes modelos.

Em suma, neste objetivo pretende-se verificar diferentes valências práticas que um programa de BI pode acrescentar na tomada de decisão de uma empresa agrícola, sem qualquer outra alteração aos seus processos produtivos.



### **3. Material e Métodos**

#### **3.1. Power BI®**

O Power BI® é um sistema de BI desenvolvido pela Microsoft Corporation®, lançado ao público em 2014, possibilitando a análise detalhada de uma empresa através da utilização de dados de diferentes formatos e fontes, incluindo fontes on-line. É uma ferramenta analítica que torna possível a interligação e atualização de um grande volume de dados, em qualquer lugar e em diversos tipos de dispositivos, permitindo a criação de modelos ou bases de dados dinâmicas, constantemente atualizáveis.

Uma das valências mais relevantes do Power BI® reside na simplicidade do processo de criação de bases de dados. O programa tem a capacidade de fazer a edição das várias consultas digitais, pelo que a grande maioria dos dados originais não necessita de alterações previamente à sua utilização. Além disso, a introdução de dados pode ser conseguida de forma completamente automatizada.

O Power BI® permite efetuar a ligação entre ficheiros e/ou bases de dados, criando uma apresentação simplificada da informação e respetiva análise. Para além disso, possibilita a elaboração de relatórios dinâmicos, que se adaptam às necessidades do utilizador final.

Uma outra valência deste programa consiste na possibilidade de partilha de relatórios dinâmicos. O acesso e edição destes últimos, por diferentes utilizadores, facultam uma maior cooperação entre diferentes departamentos de uma mesma empresa, ou mesmo entre várias empresas.

O Power BI® destaca-se, efetivamente, pela versatilidade que oferece ao utilizador, respeitante aos diferentes tipos e quantidade de dados que permite abordar, abrindo oportunidades de aplicação a áreas que não apenas a financeira. A personalização de relatórios e a flexibilidade das linguagens informáticas que usa permitem a transversalidade de utilização em empresas de diferentes áreas, incluindo as do setor agroalimentar. Permite também uma análise conjunta de diferentes departamentos numa mesma organização, através de relações entre as respetivas bases de dados, abrindo assim uma janela de oportunidade para novos indicadores mais englobantes que conferem uma visão mais abrangente do que os métodos de tratamentos de dados convencionais.

As valências enunciadas do programa são apresentadas num ambiente protegido contra fugas de informação ou ataques informáticos, também desenvolvido pela Microsoft Corporation®.

O Power BI® pode ser utilizado numa versão gratuita, com limitações na partilha de informação, existindo a opção de subscrição que desbloqueia a totalidade das funcionalidades. A versão gratuita do Power BI® integra três funcionalidades principais para a criação de modelos ou bases de dados: (1) a obtenção e edição de consultas, que consiste na importação de ficheiros para o modelo a trabalhar e a sua reestruturação para uma forma compatível com os restantes dados; (2) a criação de relações entre tabelas, assim como cálculo de novas medidas e indicadores; e, por último, (3) a apresentação

de dados em painéis dinâmicos, altamente personalizáveis em função das necessidades do utilizador final.

### 3.1.1. Obtenção e edição de consultas

Como referido, o programa Power BI® possibilita a utilização de vários tipos e formatos de ficheiros. Dada a vasta variedade de ficheiros, com denominações distintas, os ficheiros que o programa importa para a realização do processamento analítico designam-se consultas (*query*).

Os dados a utilizar, relacionados com os processos em estudo, nem sempre se encontram no mesmo ficheiro ou na mesma forma, pelo que se torna necessário realizar um tratamento estrutural, previamente às implementações de visualizações e aos cálculos analíticos.

Para a importação e uniformização de tabelas de dados é utilizado o editor de consultas. Esta ferramenta do Power BI® permite a importação e transformação de tabelas, contidas em diferentes ficheiros, para um formato adequado aos nossos objetivos de trabalho. Possui também, como uma das principais particularidades, o registo da totalidade dos passos efetuados para a transformação da tabela. Isto significa a possibilidade de introdução consecutiva de novos ficheiros, acompanhada da formatação automática das tabelas, utilizando os passos previamente registados, obtendo-se imediatamente uma tabela final com dados atualizados e dispostos da mesma forma que os anteriormente importados.

Numa primeira fase, a importação de ficheiros pode ser realizada por indicação direta da localização dos mesmos. Os ficheiros podem ser importados com diferentes formatos, como Microsoft Excel (.xlsx, .xls), Microsoft Access (.accdb), Ficheiro de Texto (.txt), JSON, entre outros. No entanto, também poderão ser importados ficheiros do mesmo formato diretamente de uma pasta ou de uma drive constituindo uma única consulta. Para tal, será necessária a configuração prévia de modo como os ficheiros devem ser fundidos. Esta última forma de importar consultas é importante porque se traduz numa simplicidade em anexar dados atualizados, bastando somente colocar os novos dados atualizados numa pasta para toda a consulta importada ser atualizada.

Esta ferramenta de edição de consultas utiliza a linguagem informática M (*Power Query M formula language*), que consiste numa linguagem informática otimizada para a construção de consultas altamente flexíveis e funcionais. A interface de utilizador apresenta uma vasta diversidade de opções que mitigam a necessidade de conhecimento profundo da linguagem M para uma formatação correta das consultas. No entanto, o utilizador desta ferramenta necessita de conhecer a linguagem M para realizar formatações mais personalizáveis e também para melhorar a sua eficiência na utilização da ferramenta.

Para um funcionamento ótimo do modelo nas fases posteriores, a edição das consultas deve resultar num formato idêntico ao formato da Tabela 2, ou seja: (1) um único cabeçalho para os diferentes atributos; (2) a edição correta do tipo de dados correspondentes a cada atributo; (3) a mínima ocorrência de valores em branco ou nulos e (4) a verificação de que para um determinado registo

somente existe um valor por atributo. Após a edição, todas as consultas devem apresentar este formato, tornando possível o trabalho analítico nas etapas posteriores sem erros relativos à forma da tabela.

Tabela 2 – Exemplo de tipologia de formatação uniforme.

	$I^2_3$ #	Data	$I^2_3$ Numero	$I^2_3$ Campanha	$I^2_3$ Operacao	$A^B_C$ Descrição	$I^2_3$ Supervisor	$A^B_C$ Nome Supervisor
1	50962	18/12/2015	9320	2015	119	Lavar Caixas	284	ANTÓNIO MIGUEL DA SILVA
2	51549	04/12/2015	9258	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
3	51550	04/12/2015	9258	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
4	51551	04/12/2015	9258	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
5	51552	04/12/2015	9258	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
6	51553	04/12/2015	9258	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
7	51554	04/12/2015	9258	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
8	51611	03/12/2015	9249	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
9	51612	03/12/2015	9249	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
10	51613	03/12/2015	9249	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
11	51675	02/12/2015	9239	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
12	51676	02/12/2015	9239	2015	10	Tirar Panos	294	MANUEL JOAQUIM LEOCÁI
13	51774	01/12/2015	9229	2015	169	Levantar Saias Definitivo	294	MANUEL JOAQUIM LEOCÁI
14	51775	01/12/2015	9229	2015	169	Levantar Saias Definitivo	294	MANUEL JOAQUIM LEOCÁI
15	51776	01/12/2015	9229	2015	169	Levantar Saias Definitivo	294	MANUEL JOAQUIM LEOCÁI
16	51777	01/12/2015	9229	2015	169	Levantar Saias Definitivo	294	MANUEL JOAQUIM LEOCÁI
17	51778	01/12/2015	9229	2015	169	Levantar Saias Definitivo	294	MANUEL JOAQUIM LEOCÁI
18	51779	01/12/2015	9229	2015	169	Levantar Saias Definitivo	294	MANUEL JOAQUIM LEOCÁI
19	51780	01/12/2015	9229	2015	169	Levantar Saias Definitivo	294	MANUEL JOAQUIM LEOCÁI
20	51781	01/12/2015	9229	2015	169	Levantar Saias Definitivo	294	MANUEL JOAQUIM LEOCÁI
21	51782	01/12/2015	9229	2015	169	Levantar Saias Definitivo	294	MANUEL JOAQUIM LEOCÁI

Para a obtenção de uma consulta formatada como a ilustrada na Tabela 2 são necessárias várias operações no editor de consultas recorrendo à linguagem informática M. Estes passos são registados, o que permite a execução dos mesmos para novos conjuntos de dados, o que significa que para dados originais com o mesmo formato entre eles, basta realizar a edição uma primeira vez. Posteriormente, para a introdução de novos grupos de dados basta incorporar os mesmos na pasta de onde os dados são importados. Esta funcionalidade confere uma facilidade acrescida da atualização do modelo por parte do utilizador final, pois não requer nenhum conhecimento da linguagem M, somente requer conhecimentos básicos de informática como a execução de uma operação de copiar e colar os dados para a pasta fonte do modelo.

### 3.1.2. Cálculos e relações entre consultas

Uma vez uniformizadas as consultas a utilizar no modelo, o Power BI® permite a realização de operações com base nas informações das tabelas de cada consulta. Estas operações podem ser realizadas numa só tabela. No entanto, o Power BI® permite também a realização de operações que abrangem várias tabelas através de relações entre atributos das mesmas. Por outra palavras, se duas tabelas distintas possuírem um atributo em comum é possível realizar operações entre todos os atributos das duas tabelas, partindo do utilizador o critério da realização dessas mesmas operações.

### 3.1.2.1. Operações

As principais operações que se podem realizar nas tabelas, dentro do modelo, são a criação de relações entre tabelas, a criação de colunas calculadas e a criação de medidas. Para a realização destas operações, a linguagem informática utilizada é o DAX (*Data Analysis Expressions*) que consiste num inventário de funções e operadores que podem ser combinados de forma a criar fórmulas e expressões em programas como o Microsoft SQL Server Analysis Services® e a extensão Power Pivot no Microsoft Excel® e também no Power BI®. A partir desta linguagem informática é possível a criação de uma grande diversidade de funções que servem vários objetivos relativos ao que o utilizador pretende. Um aspeto desta linguagem, que merece destaque nesta fase por ser relevante ao longo do trabalho desenvolvido, é a possibilidade de utilização de filtros no código criado sobre as tabelas e dados a trabalhar, o que confere ao utilizador uma maior flexibilidade no momento da elaboração de colunas calculadas ou medidas, como será demonstrado adiante.

### 3.1.2.2. Criação de relações

A criação de relações entre consultas são um grupo de operações cruciais para o trabalho de cálculo de colunas e medidas. A operação de criação de relações é relativamente simples, não exigindo muito conhecimento de qualquer linguagem informática. No entanto para que as relações sejam relevantes é necessário que o utilizador conheça e planeie o que pretende realizar nas fases posteriores de trabalho analítico.

A relação entre duas consultas é realizada através da junção de dois atributos, cada um pertencente a uma consulta. Se executada corretamente, a relação irá permitir ao utilizador o tratamento analítico utilizando informações de ambas as tabelas, permitindo assim uma síntese de informação entre as duas consultas, assim como a criação de novos indicadores integradores da informação produzida pelas duas consultas em conjunto.

Num modelo em que existam mais de duas consultas, as relações permitem a junção de informações das diferentes consultas e a criação de um contexto integrado no modelo para tratamentos analíticos posteriores. Em suma, a criação de relações permite ao Power BI® a criação de uma forma de inteligência, conhecendo como é que os diferentes elementos do modelo se condicionam, o que se traduz numa maior versatilidade e dinâmica dos indicadores criados nas etapas posteriores.

Para a criação de relações entre duas tabelas através dos seus atributos é necessário que o utilizador possua um conhecimento aprofundado sobre os dados e o seu significado nas tabelas nas quais se encontra a trabalhar e, por outro lado, que compreenda a forma como ambos os atributos se relacionam, para posteriormente conseguir executar as operações que deseje.

As relações entre duas tabelas podem ser de três tipos distintos e possuem dois tipos de direção. No caso dos tipos de relações, estes podem ser de um-para-um, de muitos-para-um ou de um-para-muitos. Isto significa que no caso de a relação ser do tipo um-para-um, os registos de um atributo

numa tabela possuem uma correspondência unívoca para os registos do atributo pertencente a outra tabela. No caso de a relação ser do tipo um-para-muitos, significa que os registos de um atributo numa tabela possuem uma correspondência não unívoca para os registos do atributo na outra tabela a relacionar. No caso do tipo muitos-para-um, acontece o oposto da relação um-para-muitos.

As direções da relação podem ser unidirecionais ou bidirecionais, o que significa que no caso de uma relação unidirecional, a relação só irá afetar uma das duas tabelas a relacionar e no caso de relações bidirecionais, ambas as tabelas serão afetadas pela relação. O efeito da relação entre as tabelas é posteriormente refletido nas visualizações das colunas calculadas e medidas criadas.

### **3.1.2.3. Colunas calculadas**

Um outro grupo de operações que podem ser realizadas nesta área do Power BI® é a criação de colunas calculadas. Este grupo de operações permite ao utilizador criar colunas numa tabela existente, tendo por base atributos da mesma tabela ou de tabelas relacionadas. A criação de colunas calculadas confere uma mais valia em termos de versatilidade no modelo, uma vez que, conjugada com a linguagem DAX, permite o acréscimo de nova informação numa tabela já existente.

### **3.1.2.4. Medidas**

Por fim, o último grupo de operações principais nesta área do programa é o cálculo de medidas. As medidas calculadas têm como característica fundamental o facto de o seu resultado ser um único valor, ou seja, para qualquer operação dentro da linguagem DAX entre um ou mais atributos, o resultado será sempre único. No entanto este apresenta a possibilidade de ser filtrado para que apresente um resultado diferente em função das seleções pretendidas.

### **3.1.3. Relatórios dinâmicos do modelo (*Dashboards*)**

Para a exposição e partilha dos resultados obtidos através das edições e criação de medidas, o Power BI® disponibiliza a criação de relatórios dinâmicos, na forma de painéis, que têm a capacidade de expor dinamicamente os resultados da análise dos dados.

Esta funcionalidade do programa permite que o utilizador crie os seus painéis de uma forma personalizada através de múltiplos visuais (gráficos, tabelas, matrizes, filtros, entre outros auxiliares visuais). Os auxiliares visuais estão relacionados com as consultas e têm a capacidade de se alterarem em função dos filtros seleccionados em cada momento de utilização.

Na criação de painéis para o modelo, o utilizador do programa tem a liberdade de eleger a melhor disposição e conteúdo visual para expor os seus resultados em função das necessidades de quem consulta os painéis.

A partilha de informação dentro de uma organização é realizada através destes painéis, sendo que todos os cálculos e edições se encontram do lado de quem criou os painéis. No entanto, quem recebe os painéis partilhados continua a possuir a capacidade de filtrar os resultados com os elementos da tabela, possibilitando assim uma consulta personalizada sem alteração do conteúdo das medidas e edição de tabelas.

Em suma, os painéis do modelo servem o propósito de relatórios dinâmicos, em que a informação exposta logra a adaptação às necessidades do utilizador em cada consulta, estando sempre atualizados. Desta forma o utilizador não precisa de adaptar a informação de relatórios estáticos às suas necessidades, assim como não necessita de atualizar manualmente os mesmos.

#### **3.1.3.1. Filtragem por segmentação de dados**

A filtragem por segmentação de dados nos painéis são uma funcionalidade presente no Power BI® que permite a seleção de atributos presentes na base de dados, alterando os apoios visuais em função da seleção, sem que seja necessária uma alteração dos dados. Ao ser selecionado um filtro, os dados dispostos e tratados no painel (ou em todo o modelo) são restringidos à seleção realizada. Isto significa que um painel pode apresentar diversos resultados em função da seleção do utilizador. Tendo em consideração que esta seleção de filtros é facilmente reversível, os filtros de segmentação de dados conferem ao painel a dinâmica de possibilitar a consulta de informação distinta para o utilizador em qualquer momento.

A funcionalidade de criação de filtros de segmentação de dados no painel é das principais operações na criação deste tipo de visualizações, uma vez que exige o conhecimento das necessidades de quem irá utilizar o mesmo. Se os filtros forem ao encontro das necessidades do utilizador final, é possível num único painel apresentar informação suficiente para satisfazer todas as necessidades do utilizador final através da dinâmica dos restantes auxiliares visuais, que são alterados pela seleção de filtros.

#### **3.1.3.2. Visuais**

Outro aspeto importante dos painéis do Power BI® é a diversidade de auxiliares visuais que este dispõe para apresentar os resultados do processamento dos dados. Esta apresentação pode ser feita através de uma grande diversidade de gráficos, mas também de tabelas ou matrizes. Para além dos auxiliares disponíveis originalmente no programa, existe um sítio na internet com ainda mais auxiliares criados pela comunidade utilizadora do programa, que vão ao encontro das necessidades

mais frequentes dos utilizadores respeitantes aos auxiliares que não se encontram disponíveis na versão original do Power BI®.

Todos os auxiliares visuais partilham a capacidade de serem segmentados, o que se traduz na capacidade de se alterarem em função das necessidades do utilizador final.

Outro aspeto importante é a capacidade de muitos auxiliares visuais serem interativos, ou seja, ao seleccionarmos alguma parte do visual, este atua como filtro em todo o painel, pelo que a dinâmica do mesmo aumenta significativamente.

### **3.1.3.3. Partilha**

Finalmente, os painéis apresentam a funcionalidade de serem partilhados entre os membros da empresa ou serem publicados na forma de endereço de internet, acessíveis a quem tenha acesso. Esta partilha contempla somente o painel e não a base de dados a que este acede, permitindo uma partilha segura sem que os dados originais se encontrem de alguma forma comprometidos.

Graças a esta funcionalidade de partilha, os painéis podem ser utilizados em qualquer lugar, numa ampla variedade de dispositivos. Permitem por exemplo, uma tomada de decisão em proximidade com o local onde essa decisão vai incidir ou mesmo a redução de distância e de tempo entre a tomada de decisão e a sua execução. Efetivamente, pode acontecer que o próprio funcionário que executa a decisão seja o mesmo que a toma, tendo em conta que a informação já se encontra tratada pelo programa, aumentando assim a eficiência de alguns processos.

Por outro lado, a consulta da situação da empresa pode ser feita em tempo real e ajustada a diferentes necessidades de informação, possibilitando assim uma visão analítica rapidamente acessível, sem ser necessário reunir ou requisitar informação presencialmente a outros membros da organização.

## **3.2. Método aplicado ao caso de estudo**

A metodologia proposta e aplicada na elaboração deste caso de estudo de criação, implementação e validação do programa Power BI® divide-se em 6 fases distintas (Figura 3):

- 1) Obtenção de dados: Realização de reuniões com responsáveis da empresa para a apresentação do programa e do caso de estudo, assim como a cedência dos primeiros dados a tratar e da identificação de alguns problemas preliminares;
- 2) Modelo de demonstração: Criação de um modelo com base nos dados fornecidos, de forma a demonstrar as capacidades do programa apresentado e validar a continuação da implementação;
- 3) Determinação de necessidades: Identificação das principais limitações na análise dos dados e da informação pretendida por cada utilizador do programa;

- 4) Modelação: Criação e implementação dos modelos, em conformidade com as respostas da fase anterior;
- 5) Verificação: Será realizada uma análise sobre os modelos criados, contando com os intervenientes nas diferentes áreas de interesse relacionadas com o programa desenvolvido;
- 6) Avaliação e discussão de resultados: Realização de uma discussão sobre os resultados deste caso de estudo.

Todo este processo e metodologia utilizada pretende conferir um contributo importante no que diz respeito à implementação de tecnologia de BI nas empresas agrícolas, em resposta ao desafio emergente da *Big Data* na agricultura. O diagrama da metodologia aplicada (Figura 3) apresenta um resumo sintético da direção desta proposta.

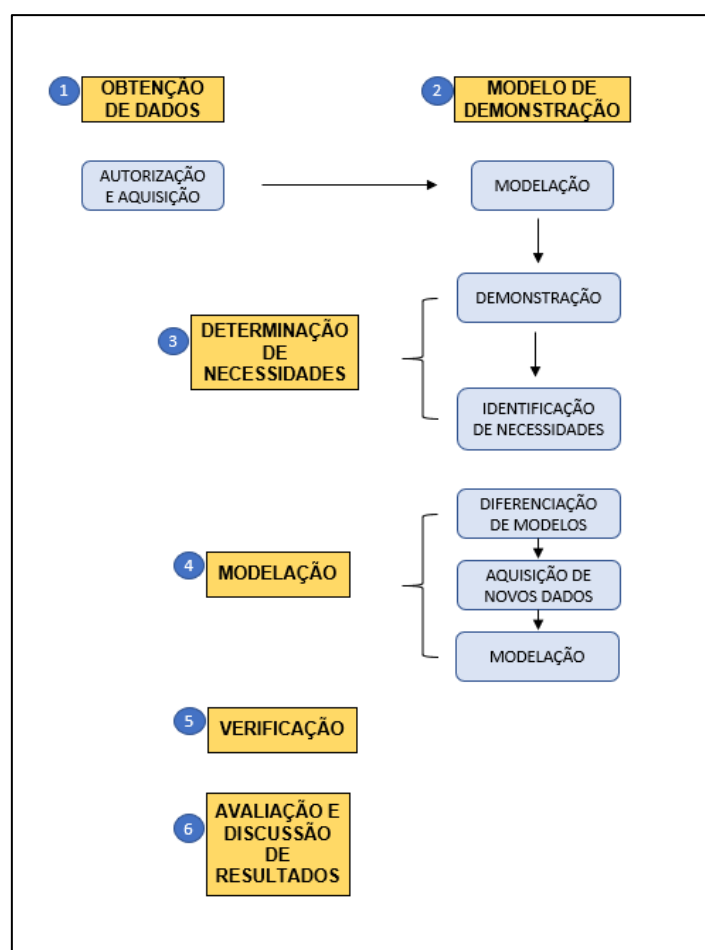


Figura 3 – Modelo conceitual do método aplicado na criação, implementação e validação da utilização do programa Power BI®.

### 3.3. Dados cedidos pela Herdade Vale da Rosa

A Herdade Vale da Rosa detém ficheiros informáticos com registos de dados respeitantes a diferentes áreas da sua produção, objetivos e planeamento de processos.



Os dados referidos são essenciais para a criação de modelos que logrem uma análise dos processos e consumos da atividade produtiva da empresa. Por esta razão foram cedidos no seu estado original, com a finalidade de incorporação em modelos que permitam a sua mais profunda análise, integrados nas necessidades de informação de quem os consulta regularmente durante a execução das suas tarefas profissionais.

As tabelas que ilustram este capítulo, correspondem às versões originais cedidas, sem nenhuma alteração quanto à sua forma ou conteúdo. No entanto, devido às dimensões das mesmas (aproximadamente 143.000 registos processados, cada um com vários atributos), estas correspondem apenas a uma amostra reduzida da sua dimensão original, servindo o propósito de contextualização.

### 3.3.1. Dados de consumos

Os dados de consumos dizem respeito à utilização de insumos na produção de uva de mesa, nomeadamente, água, adubos, correctivos, produtos fitofármacos e sementes.

Nestes dados estão incluídos os registos de utilização de insumos, assim como os programas de consumos no início de cada campanha.

#### 3.3.1.1. Registos de consumos

O ficheiro de registo de consumos corresponde a uma tabela em formato Excel que regista todos os consumos de insumos, categorizando atributos relativos a cada consumo como ilustrado na Figura 4.

SEQ	DATA	DOC	CAMPANHA	STATUS	COD_OPERACAO	DESCRICAO_OPERACAO	SUPERVISOR	NOME_SUPERVISOR	COD_ARTI	NOME_ARTIGO	QUANT	VALOR_TOTAL	DOSAGEM_HA	QUADRA	CARREIRA	AREA
21	20141103	TRAT-2600	2015	2	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	78,3428	16,84371169	100	4MB2	27	0,84
22	20141103	TRAT-2600	2015	2	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	81,1408	17,44527283	100	4MB3	28	0,87
23	20141103	TRAT-2600	2015	2	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	81,0917	17,43471912	100	4MB4	28	0,869474
24	20141103	TRAT-2600	2015	2	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	84,4051	18,14709415	100	4MB5	29	0,905
25	20141103	TRAT-2600	2015	2	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	84,3498	18,13519887	100	4MB6	29	0,904407
26	20141103	TRAT-2600	2015	2	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	40,6698	8,744003339	100	4MB7	14	0,435066
27	20141105	TRAT-2602	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	38,7629	8,334016666	100	4MB7	12	0,37377
28	20141105	TRAT-2602	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	69,1906	14,87598526	100	4MB1	26	0,66717
29	20141105	TRAT-2602	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	50,7466	10,91050997	100	2CR5	19	0,489323
30	20141105	TRAT-2602	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	50,8931	10,94201184	100	2CR6	19	0,490736
31	20141105	TRAT-2602	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	47,8289	10,28320446	100	2CR7	18	0,461189
32	20141105	TRAT-2602	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	101,091	21,73457561	100	3ME1	33	0,974769
33	20141105	TRAT-2602	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	66,926	14,38909404	100	3SW1	22	0,645333
34	20141105	TRAT-2602	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	24,5609	5,280602147	100	2SG1	9	0,236829
35	20141106	TRAT-2604	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2811	Luzerna Altiva	10	#VALOR!	5	4WS1	50	1,55303
36	20141106	TRAT-2604	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2263	Ervilhaca Libia (Vici	100	#VALOR!	50	4WS1	50	1,55303
37	20141106	TRAT-2604	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	100	21,5	50	4WS1	50	1,55303
38	20141106	TRAT-2605	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	100	21,5	50	4IT1	66	2,05
39	20141106	TRAT-2605	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2811	Luzerna Altiva	40	#VALOR!	20	4IT1	66	2,05
40	20141106	TRAT-2605	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	100	21,5	50	4WS2	66	2,05
41	20141106	TRAT-2605	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2811	Luzerna Altiva	40	#VALOR!	20	4WS2	66	2,05
42	20141110	TRAT-2607	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	9,97774	2,145214246	100	3TH1	3	0,087321
43	20141110	TRAT-2607	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	10,0241	2,155184567	100	3SG2	3	0,087727
44	20141110	TRAT-2607	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	9,99815	2,149601187	100	3SG3	3	0,0875
45	20141110	TRAT-2608	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	31,7283	6,821594811	50	3SW1	21	0,63
46	20141110	TRAT-2608	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2263	Ervilhaca Libia (Vici	31,7283	#VALOR!	50	3SW1	21	0,63
47	20141110	TRAT-2608	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	25,0286	5,381152232	50	4WS1	16	0,49697
48	20141110	TRAT-2608	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2263	Ervilhaca Libia (Vici	25,0286	#VALOR!	50	4WS1	16	0,49697
49	20141110	TRAT-2608	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	103,243	22,19725296	50	4IT2	66	2,05
50	20141110	TRAT-2608	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2263	Ervilhaca Libia (Vici	103,243	#VALOR!	50	4IT2	66	2,05
51	20141117	TRAT-2609	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	197,342	42,42860234	100	3SG2	66	1,93
52	20141117	TRAT-2609	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	178,937	38,47153062	100	3SG3	60	1,75
53	20141117	TRAT-2609	2015	1	21	ABUBAÇÃO: mecanizada	9999	Tratoristas	2640	AVEIA	166,667	35,83348281	100	3TH1	56	1,63

Figura 4 – Excerto da tabela de consumos de 2015 a 2018.

A tabela de consumos contém informação relativa às campanhas de 2015, 2016, 2017 e 2018. Cada registo possui 16 atributos que distinguem cada consumo. Detalhadamente:

- O atributo “SEQ” consiste na contagem sequencial de cada registo por ordem de entrada no ficheiro;
- O atributo “DATA” informa sobre a data de execução do consumo;
- O atributo “DOC” é um código informativo do registo de consumo, exclusivo para cada linha de registo;
- O atributo “CAMPANHA” diz respeito à campanha sobre a qual o consumo incide;
- O atributo “STATUS” confere uma ordem hierárquica ao consumo, relevante para alguns tipos de análises;
- O atributo “COD\_OPERACAO” constitui um código exclusivo para cada tipo de operação em que o consumo foi executado;
- O atributo “DESCRICAO\_OPERACAO” expõe no formato de texto o tipo de operação em que o consumo foi executado;
- O atributo “SUPERVISOR” corresponde ao código exclusivo do supervisor que realizou a operação;
- O atributo “NOME\_SUPERVISOR” encontra-se no formato de texto e diz respeito ao nome do supervisor que esteve envolvido no consumo registado;
- O atributo “COD\_ARTIGO” consiste no código exclusivo que cada tipo de artigo consumido detém;
- O atributo “NOME\_ARTIGO” representa em formato de texto a designação de cada artigo consumido;
- O atributo “QUANT” expressa a quantidade de artigo consumido;
- O atributo “VALOR\_TOTAL” consiste no valor monetário que cada consumo representou;
- O atributo “DOSAGEM\_HA” representa a dosagem de cada consumo por unidade de área, neste caso o hectare;
- O atributo “QUADRA” confere a informação da quadra onde foi realizado o consumo;
- O atributo “CARREIRA” confere a informação da carreira onde foi aplicado o consumo;
- O atributo “AREA” consiste na área em que se realizou o consumo.

Foram cedidos dois ficheiros neste formato, sendo um respeitante ao período entre janeiro de 2015 a fevereiro de 2018 e outro respeitante aos meses de março, abril e maio de 2018.

Relativamente à granularidade espacial dos registos presentes nestas tabelas de registo de consumos, os dados encontram-se ao nível da quadra (ver capítulo 2.1.).

### **3.3.1.2. Programa inicial de consumos**

Para cada campanha, a Herdade Vale da Rosa dispõe de um ficheiro de programação inicial de consumos a realizar. Foram disponibilizados os ficheiros referentes às campanhas de 2016, 2017 e

[illegible][illegible]

Nestes ficheiros estão presentes dois tipos de tabelas, uma respeitante aos consumos a executar entre as fases de abrolhamento e frutificação e outra respeitante aos consumos a executar entre as fases de engrossamento e maturação. No canto superior esquerdo do ficheiro encontra-se uma única tabela com os produtos a utilizar e a sua formulação química. Na zona inferior encontram-se as tabelas de programação de consumos, existindo uma tabela para cada variedade presente na Herdade Vale da Rosa.

33

A tabela de programação de consumos no seu estado original possui vários atributos hierarquizados, respeitantes à fase fenológica e à escala temporal e também ao tipo de produto e à sua dosagem.

Na Figura 5, que ilustra a tabela de programação de consumos desde o abrolhamento até à frutificação, as linhas correspondem ao tipo de produto a utilizar. Nas colunas, os atributos representados apresentam dois níveis hierárquicos. O atributo superior corresponde à fase do ciclo fenológico da videira e o inferior ao número de semanas decorridas após a última fase do ciclo. Os valores desta tabela correspondem à dosagem de produto em quilogramas por hectare.

Na Figura 6, em que se encontra ilustrada a tabela de programação de consumos desde a floração até à maturação, a organização dos atributos e das linhas é distinta. As linhas desta tabela correspondem à variedade programada, assim como à constituição individualizada da formulação de cada produto. Nos atributos da tabela, começando pelo atributo colocado superiormente, encontram-se as fases do ciclo fenológico da videira, o número de semanas decorridas desde o último estado fenológico e a abreviatura do produto, respetivamente. Os valores desta tabela são as dosagens a aplicar em quilogramas por hectare.

Quanto à granularidade dos registos presentes nesta tabela de planeamento inicial, a nível espacial, os dados encontram-se ao nível da variedade.

### **3.3.2. Dados de operações**

Para as operações culturais executadas na Herdade Vale da Rosa existe um registo informático que guarda individualmente cada uma e as categoriza em função de um conjunto de atributos, permitindo assim a sua distinção e posterior análise.

Neste capítulo estão presentes os registos de consumos entre os anos de 2015 e 2018 e os objetivos anuais dos mesmos.

#### **3.3.2.1. Registo de operações**

No âmbito da obtenção de dados relacionados com as operações culturais realizadas na Herdade Vale da Rosa, foram cedidos dois ficheiros Excel que consistem numa tabela onde se encontra o registo de todas as tarefas culturais realizadas entre 2015 e 2018, definindo um conjunto de atributos que individualizam e categorizam cada tarefa realizada.

Na análise deste registo existem dois níveis de granularidade das atividades, a operação e a tarefa, sendo que uma operação pode integrar várias tarefas.

Na Figura 7 está ilustrada parte da tabela de registos de operações em que cada registo está categorizado através de 13 atributos distintos.

#	Data	Numero	Campanha	Operacao	Descricao	Supervisor	Nome Supervisor	Funcionarios	HorasFunc	Quadra	Carreiras	Hectares
1	20180510	3902	2018	45	Destroçar	302	JOÃO DIAS GUERF	2	1,02	1RG12	12	0,29
2	20180510	3902	2018	45	Destroçar	302	JOÃO DIAS GUERF	2	3,49	1V12	41	1
3	20180510	3902	2018	45	Destroçar	302	JOÃO DIAS GUERF	2	3,49	1V18	41	1
4	20180510	3904	2018	148	Manutenç	302	JOÃO DIAS GUERF	2	16	0GEN	1	1
5	20180510	3906	2018	25	Rega - Ma	302	JOÃO DIAS GUERF	2	3,63	4S1	61	1,73
6	20180510	3906	2018	25	Rega - Ma	302	JOÃO DIAS GUERF	2	2,85	4S2	48	1,36
7	20180510	3906	2018	25	Rega - Ma	302	JOÃO DIAS GUERF	2	2,18	4S3	33	1,04
8	20180510	3906	2018	25	Rega - Ma	302	JOÃO DIAS GUERF	2	2,24	4S4	34	1,07
9	20180510	3906	2018	25	Rega - Ma	302	JOÃO DIAS GUERF	2	1,78	4S5	31	0,85
10	20180510	3906	2018	25	Rega - Ma	302	JOÃO DIAS GUERF	2	1,68	4S6	23	0,8
11	20180510	3906	2018	25	Rega - Ma	302	JOÃO DIAS GUERF	2	1,64	4S7	30	0,78
12	20180510	3928	2018	18	Amarrar L	3804	RICARDO ANDRE	13	30,22	1P11	41	1
13	20180510	3928	2018	18	Amarrar L	3804	RICARDO ANDRE	13	44,34	1P12	61	1,47

Figura 7 - Excerto da tabela de registo de operações de 2015 a 2017.

Concretamente:

- O atributo “#” consiste na ordem sequencial de entrada dos registos no ficheiro;
- O atributo “Data” consiste na data em que cada tarefa foi realizada;
- O atributo “Numero” diz respeito ao código de ordem de execução de cada tarefa;
- O atributo “Campanha” detalha a campanha em que a tarefa foi realizada;
- O atributo “Operação” é um código exclusivo de identificação do tipo de operação realizada;
- O atributo “Descrição” consiste na descrição na forma de texto da tarefa realizada;
- O atributo “Supervisor” corresponde ao código exclusivo de cada supervisor responsável por cada tarefa;
- O atributo “Nome Supervisor” atribui a cada registo de operação o nome do supervisor alocado a cada tarefa registada;
- O atributo “Funcionarios” refere o número de funcionários alocados a cada tarefa;
- O atributo “HorasFunc” consiste no número de horas totais realizadas para executar a tarefa.
- O atributo “Quadra” revela a quadra onde a tarefa foi realizada;
- O atributo “Carreiras” revela a carreira onde a tarefa decorreu;
- O atributo “Hectares” consiste no número de hectares em que a tarefa se realizou.

Foram cedidos dois ficheiros neste formato, sendo que o primeiro consistiu no registo de operações nas campanhas de 2015, 2016 e 2017, enquanto o segundo disse respeito às operações da campanha de 2018 até ao mês de maio.

Quanto à granularidade dos registos presentes nestas tabelas de registo de operações, a nível espacial, os dados encontram-se ao nível da quadra. Ao nível da atividade a granularidade encontra-se ao nível da tarefa.

### 3.3.2.2 Objetivos de operações

A respeito dos dados de operações, a Herdade Vale da Rosa disponibilizou os objetivos referentes à campanha 2018 para todas as suas operações. Estes objetivos (Coeficiente de Eficiência Ajustado) são baseados no registo histórico do último ano (Figura 8). Os objetivos para 2018 em horas foram determinados pela Herdade Vale da Rosa.

Codigo/Operação	Coeficiente de Eficiencia Ajustado
A.1 - MARCAÇÃO DA VINHA	
A.2 - PLANTAÇÃO	61,8%
A.3 - MONTAGEM VINHA	60,0%
A.4 - MONTAGEM DA REGA	80,0%
A.5 - INSTALAÇÃO REDE	
A.6 - INSTALAÇÃO PLASTICO	
A.7 - MONTAGEM DRENAGEM	
B.1 - ENXERTAR	15,0%
B.2 - PODA DE FORMAÇÃO	75,0%
B.3 - AMARRAR CEPA	85,0%
C.1 - MARCAÇÃO	92,0%
C.2 - PODA SECA	98,5%
C.3 - DESTROÇAR VARAS	99,0%
C.4 - ATAR VARAS	98,5%

Figura 8 – Exemplo da tabela de coeficientes de objetivos de operações para a campanha 2018

A tabela de objetivos para 2018 (Figura 8) é composta por 2 atributos, respetivamente:

- O atributo “Codigo/Operação”, que informa sobre a operação a que cada registo corresponde;
- O atributo “Coeficiente de Eficiencia Ajustado”, que corresponde a um coeficiente entre o registo das horas utilizadas no ano anterior e do objetivo de horas a utilizar para a campanha de 2018. Concretamente, se para uma determinada operação, o objetivo de utilização de horas para 2018 for superior à utilização de horas no ano anterior, o coeficiente irá ser superior a 100%.

A granularidade destes objetivos encontra-se ao nível da operação. O coeficiente de eficiência ajustado mantém-se inalterado para as tarefas incluídas em cada operação.

### 3.3.3. Dados de colheita

A Herdade Vale da Rosa disponibilizou os registos de colheita e os seus objetivos para o período entre 2016 e 2018. Estes registos foram obtidos através de registos maioritariamente

analógicos. Na campanha de 2018 a empresa encontra-se em processo de instalação de controlo de colheita digital, através de *tablets* com ligação à internet e leitor de código de barras, que transmitem os dados de colheita em tempo real para a base de dados da empresa, o que significa que os dados de 2018 presentes nesta tabela foram já incorporados automaticamente.

### 3.3.3.1. Registo de colheita

Para o estudo dos valores registados relacionados exclusivamente com a operação cultural de colheita, foi disponibilizado um ficheiro de dados correspondentes à colheita na campanha de 2017 (Figura 9). Neste ficheiro encontram-se registadas as operações de colheita, categorizadas por alguns atributos dirigidos para uma maior relevância na sua análise.

#	Data	Artigo	Descricao	Quadra	Lote	Exploracao	Caixa	PesoLiquido	Tara	Nº Paleta Interr	Nº Superviso	Nome Supervis	HoraLancament	OrdemColheita
1	20170626	WHI64180CK	Uva Branca Whi4W52		1WS170626ASF		20	202,9	59,6	135002	3634	MARCELO ALEXANDRE C	13:03:05	9317
2	20170626	WHI64180CK	Uva Branca Whi4W52		1WS170626ASF		20	210,9	59,6	135005	3634	MARCELO ALEXANDRE C	13:15:42	9317
3	20170627	WHI64180CK	Uva Branca Whi4W51		1WS170627ASF		20	207,9	59,6	135019	3634	MARCELO ALEXANDRE C	10:46:25	9318
4	20170627	WHI64180CK	Uva Branca Whi4W52		1WS170627ASF		10	100,2	41,3	135031	3634	MARCELO ALEXANDRE C	15:43:25	9318
5	20170703	WHI64180CK	Uva Branca Whi4W51		1WS170703ASF		20	207,4	59,6	135041	3634	MARCELO ALEXANDRE C	11:39:40	9319
6	20170703	WHI64180CK	Uva Branca Whi4W51		1WS170703ASF		20	206,9	59,6	135048	3634	MARCELO ALEXANDRE C	14:25:04	9319
7	20170703	WHI64180CK	Uva Branca Whi4W51		1WS170703ASF		20	202,4	59,6	135050	3634	MARCELO ALEXANDRE C	15:30:33	9319
8	20170703	WHI64180CK	Uva Branca Whi4W52		1WS170703ASF		20	200,9	59,6	135040	2783	SARA ALEXANDRA GOMI	11:37:02	9319
9	20170704	CAR64180CK	Uva Preta Cardin1C5		1CA170704ASF		20	187,4	59,6	135080	3026	LUIS MANUEL DA SILVA I	10:51:25	9321
10	20170704	CAR64180CK	Uva Preta Cardin1C5		1CA170704ASF		11	107,87	43,13	135094	3026	LUIS MANUEL DA SILVA I	11:31:19	9321
11	20170704	WHI64180CK	Uva Branca Whi4W53 U		2WS170704UVAL		24	256,08	66,92	135134	1282	JOÃO MIGUEL DA SILVA	14:44:51	9322
12	20170704	WHI64180CK	Uva Branca Whi4W51		1WS170704ASF		20	206,4	60,6	135109	3634	MARCELO ALEXANDRE C	13:15:12	9320
13	20170704	WHI64180CK	Uva Branca Whi4W52		1WS170704ASF		20	204,9	59,6	135099	2783	SARA ALEXANDRA GOMI	11:48:10	9320
14	20170704	WHI64180CK	Uva Branca Whi4W52		1WS170704ASF		20	199,9	59,6	135116	2783	SARA ALEXANDRA GOMI	13:27:54	9320
15	20170705	CAR64180CK	Uva Preta Cardin1C17		1CA170705ASF		23	199,91	65,09	135068	3026	LUIS MANUEL DA SILVA I	10:17:26	9321
16	20170705	CAR64180CK	Uva Preta Cardin1C18		1CA170705ASF		16	160,72	52,28	135091	3026	LUIS MANUEL DA SILVA I	11:09:15	9321
17	20170705	WHI64180CK	Uva Branca Whi4W52		1WS170705ASF		20	202,4	59,6	135084	3634	MARCELO ALEXANDRE C	10:55:11	9320
18	20170705	WHI64180CK	Uva Branca Whi4W52		1WS170705ASF		20	195,9	59,6	135096	3634	MARCELO ALEXANDRE C	11:34:29	9320
19	20170705	WHI64180CK	Uva Branca Whi4W52		1WS170705ASF		20	201,4	59,6	135142	3634	MARCELO ALEXANDRE C	15:29:35	9320
20	20170706	CAR64180CK	Uva Preta Cardin1C20		1CA170706ASF		20	186,4	60,6	135150	3026	LUIS MANUEL DA SILVA I	09:28:23	9321
61	20170704	WHI64180CK	Uva Branca Whi4W51 U		2WS170704UVAL		24	253,08	66,92	135100	1282	JOÃO MIGUEL DA SILVA	11:56:53	9322
62	20170704	WHI64180CK	Uva Branca Whi4W51 U		2WS170704UVAL		24	262,08	66,92	135125	1282	JOÃO MIGUEL DA SILVA	14:07:29	9322
23	20170706	WHI64180CK	Uva Branca Whi4W51		1WS170706ASF		20	201,4	59,6	135174	3634	MARCELO ALEXANDRE C	14:35:30	9320
24	20170706	WHI64180CK	Uva Branca Whi4W51		1WS170706ASF		20	201,9	59,6	135176	3634	MARCELO ALEXANDRE C	15:18:02	9320
25	20170706	WHI64180CK	Uva Branca Whi4W51		1WS170706ASF		20	212,9	59,6	135182	3634	MARCELO ALEXANDRE C	16:31:14	9320

Figura 9 - Exemplo de tabela de registo da colheita da campanha 2017.

Na tabela ilustrada, cada registo está categorizado por 15 atributos distintos. Concretamente:

- O atributo “Exploracao” diz respeito à sigla de exploração que realizou a colheita, uma vez que existem duas explorações que realizam esta operação, a UVAL e a ASF;
- O atributo “#” diz respeito à ordem sequencial da entrada dos registos no ficheiro;
- O atributo “Data” consiste na data em que a colheita foi realizada;
- O atributo “Artigo” representa a designação interna dos produtos colhidos;
- O atributo “Descrição” representa a descrição do produto obtido em cada registo;
- O atributo “Quadra” consiste na quadra em que se realizou a colheita;
- O atributo “Lote” representa o lote de colheita no qual cada registo está incluído;
- O atributo “Caixas” consiste no número de caixas de uva obtidas em cada um dos registos de colheita;
- O atributo “PesoLíquido” é o peso das uvas colhidas em cada registo;
- O atributo “Tara” consiste no peso do material de armazenamento da uva utilizado em cada registo;
- O atributo “Nº Paleta Interna” consiste num código exclusivo para cada paleta utilizada para o processo logístico de transporte da uva colhida;
- O atributo “Nº Supervisor” consiste no número exclusivo que cada supervisor possui;

- O atributo “Nome Supervisor” confere em forma de texto o nome do supervisor alocado a cada registo de colheita;
- O atributo “HoraLançamento” corresponde à hora em que o registo de colheita foi concluído no campo;
- O atributo “OrdemColheita” consiste no código da ordem de colheita a que o registo corresponde.

Quanto à granularidade dos registos presentes nesta tabela de registo das operações de colheita, a nível espacial, os dados encontram-se ao nível da quadra.

### 3.3.3.2. Objetivos de colheita

A respeito dos objetivos de colheita, foram cedidas pela Herdade do Vale da Rosa as previsões de produção para a campanha. Na Figura 10 existem três níveis de granularidade distintos, designadamente o nível da vinha, da variedade e da quadra, cada um destes com a previsão de produtividade associada.

QUADRAS nova nomenclatura	QUADRAS antiga nomenclatura	VINHA	INVESTIMENTO X PRODUÇÃO	PRODUÇÃO 2017-2018		VARIEDADE
				PREVISÃO HECTARE	PREVISÃO TOTAL	
0GEN	0GEN		Producao			Quadra Genérica
1C1	C1	1	Producao	23 520	23 545,56	Cardinal
1C2	C2	1	Producao	25 841	25 835,60	Cardinal
1C3	C3	1	Producao	26 488	26 483,17	Cardinal
1C4	C4	1	Producao	48 934	48 953,23	Cardinal
1C5	C5	1	Producao	36 310	36 248,16	Cardinal
1C6	C6					
1C7	C7	1	Producao	43 367	43 427,31	Cardinal
1C8	C8	1	Producao	22 000	21 974,09	Cardinal
1C9	C9	1	Producao	40 058	40 002,16	Cardinal
1C10	C10	1	Producao	25 793	25 780,20	Cardinal
1C11	C11	1	Producao	41 192	41 055,77	Cardinal
1C12	C12	1	Producao	32 424	28 925,71	Cardinal
1C13	C13	1	Producao	27 028	8 713,86	Cardinal

Figura 10 – Excerto da tabela de previsão de produção para a campanha 2017-2018.

Esta tabela tem como função a síntese de previsões de produção ao nível de granularidade da quadra e é composta por 7 atributos distintos que nos permitem analisar os dados a um nível de granularidade mais elevado, assim como analisar os valores previstos. Concretamente:

- O atributo “QUADRAS nova nomenclatura” diz respeito ao código atualmente utilizado na identificação de quadras, em que o primeiro algarismo significa a vinha em que se localiza, o segundo carácter é a inicial da variedade de videira que se encontra instalada nessa quadra e o terceiro e quarto algarismos enumeram a quantidade de quadras situadas na mesma vinha



e com a mesma variedade, sendo que também cumprem uma ordem geográfica organizacional.;

- O atributo “QUADRAS antiga nomenclatura” consiste no código utilizado anteriormente à implementação do novo código, sendo necessário uma vez que se encontra em alguns registos mais antigos;
- O atributo “VINHA” corresponde à vinha onde a quadra se encontra;
- O atributo “INVESTIMENTO X PRODUÇÃO” classifica as quadras entre produção de uvas de mesa ou fase de investimento, ou seja, fase não produtiva;
- O atributo “PREVISAO HECTARE” consiste na produtividade prevista em quilogramas por hectare;
- O atributo “PREVISÃO TOTAL” diz respeito à previsão total de cada quadra em quilogramas;
- O atributo “VARIEDADE” diz respeito à variedade que se encontra instalada em cada quadra.

### 3.3.4. Dados culturais

Neste capítulo são abordados os dados técnicos culturais cedidos pela Herdade Vale da Rosa. Estes dados (Figura 11) dizem respeito ao acompanhamento do ciclo fenológico das viedas das diferentes variedades presentes na exploração.

		Estados fenológicos			
		Abrolhamento	Floração	Pintor	Maturação
Variedades	Cobertura	2017	2017	2017	2017
Cardinal	Coberta Vinha 1	08/mar	19/abr	29/mai	28/jun
	Coberta 3C1	15/mar	26/abr	13/jun	05/jul
	Descoberta	24/mar	10/mai	22/jun	12/jul
Victória	Coberta	22/mar	20/abr	15/jun	12/jul
Red Globe	Coberta	20/mar	26/abr	05/jul	23/ago
	Coberta vinha 2	17/mar	26/abr	05/jul	23/ago
	Descoberta	22/mar	05/mai	05/jul	17/ago
Palieri	Coberta	20/mar	26/abr	28/jun	01/ago
	Descoberta	22/mar	10/mai	05/jul	10/ago
Crimson	Coberta	15/mar	08/mai	05/jul	10/ago
	Descoberta	15/mar	10/mai	05/jul	10/ago

Figura 11 – Exemplo de tabela de data de ocorrência de estados fenológicos em diferentes variedades de videira, na campanha de 2017.

A tabela cedida possui a descrição das datas em que cada variedade atingiu os diferentes estados fenológicos relevantes, na campanha de 2017. Na tabela existem 6 atributos distintos em função dos registos de cada variedade e da utilização de cobertura. Mais especificamente:

- O atributo “Variedades” consiste na designação de variedades presentes na exploração;
- O atributo “Cobertura” indica se existe cobertura em cada registo;

- O atributo “Abrolhamento 2017” indica a data de ocorrência do estado de abrolhamento nos diferentes registos, assim como os atributos “Floração 2017”, “Pintor 2017” e “Maturação 2017” indicam respetivamente a data de ocorrência dos estados de floração, pintor e maturação, respetivamente.

Nestes dados, o nível de granularidade que se encontra disponível para tratamento de dados posterior encontra-se ao nível da variedade.

### 3.3.5. Dados físicos

Relativamente aos dados de cariz físico, foram cedidos parâmetros relacionados com as dimensões espaciais de cada quadra, com todos os níveis de granularidade disponíveis, assim como informações relativas ao conteúdo de cada quadra (Figura 12).

QUADRAS nova nomenclatura	QUADRAS antiga nomenclatura	U V A S			VINHA
		VARIETADE	ANO ENXERTI A	ÁREA (Ha)	
0GEN	0GEN	Quadra Genérica			
1C1	C1	CARDINALI (Vinha 1)	1997	1,00110	1
1C2	C2	CARDINALI (Vinha 1)	1997	0,99980	1
1C3	C3	CARDINALI (Vinha 1)	1997	0,99980	1
1C4	C4	CARDINALI (Vinha 1)	1997	1,00040	1
1C5	C5	CARDINALI (Vinha 1)	1997	0,99830	1
1C6	C6				
1C7	C7	CARDINALI (Vinha 1)	1997	1,00140	1
1C8	C8	CARDINALI (Vinha 1)	1997	0,99880	1
1C9	C9	CARDINALI (Vinha 1)	1997	0,99860	1
1C10	C10	CARDINALI (Vinha 1)	1997	0,99950	1
1C11	C11	CARDINALI (Vinha 1)	1997	0,99670	1
1C12	C12	CARDINALI (Vinha 1)	1997	0,89210	1
1C13	C13	CARDINALI (Vinha 1)	1997	0,32240	1

Figura 12 - Exemplo da tabela original com as características físicas das quadras.

Existem 7 atributos distintos em função de cada registo de quadra existente na Herdade Vale da Rosa. Especificamente:

- O atributo “QUADRAS nova nomenclatura” diz respeito ao código atualmente utilizado na identificação de quadras em que o primeiro algarismo significa a vinha em que se localiza, o segundo carácter é a inicial da variedade de videira que se encontra instalada nessa quadra e o terceiro e quarto algarismos enumeram a quantidade de quadras situadas na mesma vinha e com a mesma variedade, sendo que também cumprem uma ordem geográfica organizacional;
- O atributo “QUADRAS antiga nomenclatura” consiste no código utilizado anteriormente à implementação do novo código, sendo necessário uma vez que se encontra em alguns registos mais antigos;

- O atributo “VARIEDADE” indica a variedade presente em cada quadra;
- O atributo “ANO ENXERTIA” revela o ano em que as vinhas foram enxertadas;
- O atributo “ÁREA (Ha)” consiste na área de cada quadra, em hectares;
- O atributo “Vinha” representa a vinha a que cada quadra pertence.

### 3.3.6. Dados climáticos

A Herdade Vale da Rosa dispõe de uma estação meteorológica, de forma a monitorizar alguns indicadores relacionados com o clima à escala da exploração.

Os dados apresentados na Figura 13 representam parte dos dados cedidos pela Herdade Vale da Rosa que, na sua totalidade, constituem os dados climáticos referentes a um intervalo temporal desde o ano de 2015 a finais de junho de 2018.

Date	Precipitation [mm]	HC Air temperature [°C]			HC Relative humidity [%]	ETo [mm]
	Somatório	Med.	mínimo	máximo	Med.	Med.
2015-09-27 00:00:00	0	21,3	11,3	33,2	47,6	2,5
2015-09-28 00:00:00	0	19,7	12,2	27,9	66,8	3
2015-09-29 00:00:00	0	20,1	12,2	29,3	64,5	3
2015-09-30 00:00:00	0	19	10	29,4	70,7	2,3
2015-10-01 00:00:00	0	19,4	10,5	30,5	70,8	2,4
2015-10-02 00:00:00	0	19,3	10,4	30,9	72,4	2,5
2015-10-03 00:00:00	0	19,5	14,5	25,8	78,7	2,1
2015-10-04 00:00:00	1	23,2	18,9	28,7	72,8	3
2015-10-05 00:00:00	1,6	21,8	19	27,2	85,4	3,2
2015-10-06 00:00:00	0	18,9	11,8	25,3	81,1	2,3
2015-10-07 00:00:00	0	15,4	8,8	23,9	74,7	2,2
2015-10-08 00:00:00	0	14,9	5,2	26,1	76,2	2,1
2015-10-09 00:00:00	0	15,7	7,3	25,6	76	2,6
2015-10-10 00:00:00	1,2	17,5	15,3	20,1	93,2	1,6
2015-10-11 00:00:00	0,6	18,8	15,3	23,4	87,2	2,7
2015-10-12 00:00:00	10,2	18,1	16,8	21,4	98,5	1,2
2015-10-13 00:00:00	1,4	19,2	16	24,5	59,3	2,7
2015-10-14 00:00:00	0,2	17,8	12,7	25,9	72,5	1,7
2015-10-15 00:00:00	0,2	18,1	11,9	26,8	29,5	2,1
2015-10-16 00:00:00	0,2	18,8	14,8	25,5	34,9	2,2

Figura 13 - Exemplo da tabela dos dados climáticos originária da estação meteorológica da Herdade Vale da Rosa.

Nesta tabela originária da estação meteorológica da exploração estão presentes 7 atributos distintos, em função de cada registo diário. Os atributos são:

- O atributo “Data” que revela a data em que os registos foram executados;
- O atributo “Precipitação [mm] Somatório” que consiste no somatório diário de precipitação em milímetros;
- O atributo “HC Air temperature [°C] Med.”, “mínimo” e “máximo” que traduzem o registo da temperatura do ar média, máxima e mínima, respetivamente;
- O atributo “HC Relative humidity [%] Med.” que consiste na humidade relativa média;
- O atributo “ET0 [mm] med” que corresponde à evapotranspiração de referência.

### 3.3.7. Dados de qualidade

Relativamente aos dados de qualidade disponibilizados pela empresa, as tabelas possuem 54 atributos, pelo que a Figura 14 apresenta somente os atributos utilizados na modelação.

DATA	VARIEDADE	QUADRA	BRIX	CONFORMIDADE BRIX	CALIBRE	CONFORMIDADE CALIBRE	BRIX MÍNIMO PRETENDIDO	CALIBRE MÍNIMO PRETENDIDO
04-07-2016	white_seedless		17,9	C	17,0	C	17	17
04-07-2016	white_seedless		21,0	C	16,0	C	17	17
04-07-2016	white_seedless		20,1	C	17,0	C	17	17
04-07-2016	white_seedless		23,0	C	17,0	C	17	17
04-07-2016	white_seedless		23,7	C	17,0	C	17	17
04-07-2016	white_seedless		20,9	C	17,0	C	17	17
04-07-2016	white_seedless		18,6	C	17,0	C	17	17
04-07-2016	white_seedless		18,5	C	17,0	C	17	17
04-07-2016	white_seedless		20,4	C	17,0	C	17	17
04-07-2016	white_seedless		18,7	C	18,0	C	17	17

Figura 14 - Exemplo da tabela (reduzida) do controlo de qualidade.

Na tabela da Figura 14 estão presentes 9 atributos distintos, respetivamente:

- O atributo “DATA” que corresponde à data de execução de cada registo de qualidade;
- O atributo “VARIEDADE” que consiste na variedade registada;
- O atributo “QUADRA” que nos informa da quadra de origem do registo;
- O atributo “BRIX” que corresponde ao °Brix de cada registo
- O atributo “CONFORMIDADE BRIX” que consiste na comparação do °Brix da amostra registada com o “BRIX MÍNIMO PRETENDIDO”;
- O atributo “CALIBRE” que corresponde ao calibre da uva, expresso em mm;
- O atributo “CONFORMIDADE CALIBRE” que consiste na comparação do calibre registado com o “CALIBRE MÍNIMO PRETENDIDO”;
- O atributo “BRIX MÍNIMO PRETENDIDO” que consiste no °Brix mínimo pretendido para cada registo;
- O atributo “CALIBRE MÍNIMO PRETENDIDO” que corresponde ao calibre mínimo pretendido para cada registo.

## **4. Resultados da criação de modelos de BI**

A criação de modelos neste caso de estudo almeja a criação de painéis que logrem mitigar as limitações reais dos responsáveis pela tomada de decisão, no que diz respeito à análise de informação disponível. Para esse efeito, neste trabalho, a modelação segue uma metodologia em que face às diferentes versões dos modelos criados, as respostas dos utilizadores finais serão as principais fontes de alteração e otimização dos mesmos.

O objetivo fundamental dos painéis criados é, por conseguinte, suplantar as limitações de análise de dados dos utilizadores finais. Os painéis não são criados em função da quantidade máxima de indicadores que possam ser elaborados, mas sim em função de indicadores que os utilizadores considerem essenciais para a execução das suas tomadas de decisão. Um painel perde utilidade se abranger todos os indicadores teoricamente importantes, em detrimento das reais necessidades dos seus utilizadores.

O método de desenvolvimento de modelos finais passa pela criação de modelos preliminares o mais englobantes possível, tendo em consideração os dados disponibilizados e a sua apresentação aos utilizadores finais. Posteriormente, em função das respostas obtidas, os modelos finais são criados de uma forma mais focada nas necessidades dos utilizadores finais. Esta metodologia pretende isolar os indicadores, relevantes para a tomada de decisão, de entre todos os que não sejam úteis para cada utilizador em particular.

Todos os dados, indicadores e produtos presentes nos painéis dos modelos foram alterados de forma a não comprometer a confidencialidade dos mesmos. No entanto, a disposição e ilustração da funcionalidade dos painéis encontram-se inalteradas.

### **4.1. Modelo de demonstração**

Para a identificação de fatores limitantes na análise dos dados disponíveis na empresa, é fundamental que os responsáveis pela tomada de decisão experienciem uma demonstração das funcionalidades das ferramentas utilizadas para o processamento e análise dos dados fornecidos. Para isso, foi criado um modelo de demonstração que engloba todos os dados fornecidos preliminarmente e que permite a demonstração da dinâmica de vários tipos de indicadores em painéis do Power BI®.

O objetivo deste modelo de demonstração é a exposição das potencialidades do programa na análise de dados da exploração, elaborando alguns indicadores e criando relações entre dados que elucidem os utilizadores sobre os novos tipos de análise que poderá ser realizada. Para além disso, este modelo visa também a obtenção de críticas construtivas para modelos posteriores, por parte dos potenciais utilizadores, de forma a possibilitar uma aproximação às necessidades reais de informação.

Nesta fase, é também importante a verificação da utilidade das funções do Power BI® para o aumento da eficiência da tomada de decisão por parte dos elementos da empresa, assegurando a continuidade do projeto de desenvolvimento de modelos.

#### 4.1.1. Planeamento do modelo de demonstração

Nesta primeira fase, a Herdade Vale da Rosa cedeu para o modelo de demonstração dados relativos aos consumos (Figura 4), às operações (Figura 7), aos dados de colheita (Figura 9), às previsões de produção (Figura 10) e aos parâmetros físicos da exploração (Figura 12).

Este modelo de demonstração teve como finalidade ser apresentado a dois elementos da Herdade Vale da Rosa, designadamente ao responsável do departamento técnico e de qualidade e ao responsável pelo acompanhamento e análise dos dados relativos às operações na empresa, tendo um forte impacto na alocação de recursos humanos nas várias operações. O foco da demonstração seria a vertente de consumos para o primeiro e a vertente das operações para o último.

Tendo em consideração os utilizadores para os quais este modelo de demonstração foi desenvolvido e os dados disponíveis, foi planeada a criação de cinco painéis distintos com a incorporação de todos os dados enunciados.

Cada painel serve um único propósito de informação, conferindo especial relevância à comparação de valores e mudança de granularidade dos dados em tempo real.

O primeiro painel de demonstração planeado foi a “Previsão vs. Produção 2017”, que compara os dados de previsão com os dados de colheita, incorporando as granularidades espaciais da quadra. O objetivo deste painel é informar o utilizador dos desvios entre a previsão e a produção à escala da empresa, mas, no entanto, permitindo a análise e comparação desta diferença a nível da vinha, variedade e quadra. Este painel permite a rápida identificação das diferenças de produtividade e está preparado para receber dados de outros anos sem ser necessária nova modelação.

O planeamento do segundo painel de demonstração consistiu na estruturação do consumo de água na exploração. Este painel é obtido através dos dados de consumos da exploração. Uma vez mais, a granularidade espacial é importante neste modelo, permitindo a comparação entre quadras, variedades, vinhas e a totalidade da exploração, em diferentes campanhas, recorrendo a gráficos para a fácil comparação entre campanhas e granularidades físicas.

O terceiro painel planeado diz respeito à relação entre a colheita e os supervisores de equipas de colheita. Tem como objetivo mostrar a quantidade de uva colhida ao longo da campanha. Estes resultados são encontrados no painel em função da tipologia de colheita (colhido a granel ou embalado no campo) e em função da exploração que realizou a colheita uma vez que existem duas explorações que realizam esta operação, a UVAL e a ASF.

O quarto e o quinto painel deste modelo consistem em registos históricos das operações e dos consumos. Estes painéis possibilitam a consulta de cada artigo consumido ou operação realizada, individualmente, exibindo dados relativos a cada seleção em tempo real. Estes dois painéis servem o propósito de mostrar a eficiência e rapidez de uma consulta de registo histórico que contrasta com o tempo anteriormente necessário para a obtenção da mesma informação.

#### 4.1.2. Modelação

Como referido, para a elaboração deste modelo de demonstração foram importados os dados relativos a consumos, operações, colheita, previsões de produção e parâmetros físicos da exploração.

De seguida foi realizada a edição dos dados que consistiu na uniformização do formato das tabelas e dos tipos de dados para a forma da Tabela 1, através da linguagem informática M, no editor de consultas do Power BI®.

Posteriormente, foi necessária a criação de tabelas auxiliares com a finalidade de completar o modelo e facilitar o seu tratamento analítico nas etapas seguintes.

Neste âmbito foi criada uma tabela auxiliar de calendário (Tabela 3), que lista todas as datas desde o ano 2014 ao ano 2018, com alguns atributos relativos à data. A criação da tabela auxiliar calendário é transversal a todos os modelos, uma vez que a relação com os períodos de tempo é imprescindível nas diferentes análises realizadas neste caso de estudo.

Para lograr a construção de análises cruzando os dados das diferentes consultas foi necessário a criação de relações entre as mesmas. O mapa das relações criadas está ilustrado na Figura 15, em que no centro se encontram as consultas principais, que neste caso dizem respeito aos consumos, operações e à colheita. Nas extremidades estão presentes as tabelas auxiliares que correspondem à previsão de produção, ao calendário e aos parâmetros físicos da exploração.

Data	Year	Month	Month Name	Day of Week	Quarter	Week of Month	Day
05/01/2014	2014	1	janeiro	0	1	2	5
06/01/2014	2014	1	janeiro	1	1	2	6
07/01/2014	2014	1	janeiro	2	1	2	7
08/01/2014	2014	1	janeiro	3	1	2	8
09/01/2014	2014	1	janeiro	4	1	2	9
10/01/2014	2014	1	janeiro	5	1	2	10
11/01/2014	2014	1	janeiro	6	1	2	11
02/02/2014	2014	2	fevereiro	0	1	2	2
03/02/2014	2014	2	fevereiro	1	1	2	3
04/02/2014	2014	2	fevereiro	2	1	2	4
05/02/2014	2014	2	fevereiro	3	1	2	5
06/02/2014	2014	2	fevereiro	4	1	2	6
07/02/2014	2014	2	fevereiro	5	1	2	7
08/02/2014	2014	2	fevereiro	6	1	2	8
02/03/2014	2014	3	março	0	1	2	2
03/03/2014	2014	3	março	1	1	2	3
04/03/2014	2014	3	março	2	1	2	4
05/03/2014	2014	3	março	3	1	2	5
06/03/2014	2014	3	março	4	1	2	6
07/03/2014	2014	3	março	5	1	2	7
08/03/2014	2014	3	março	6	1	2	8

Figura 15 - Amostra da tabela auxiliar de calendário.

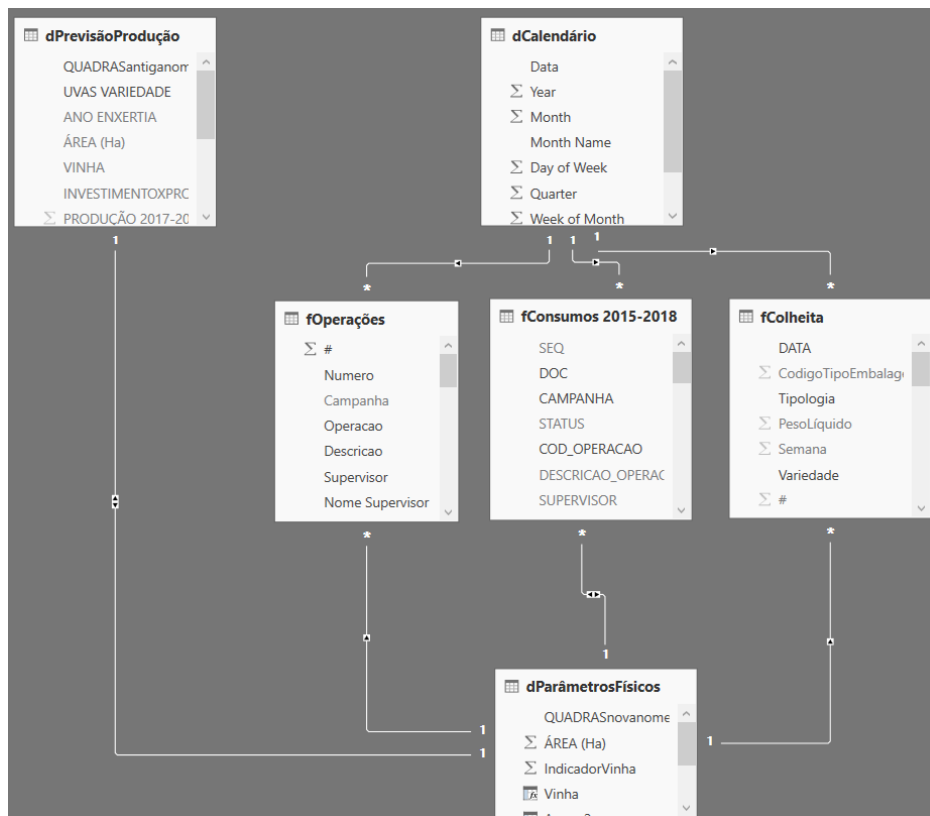


Figura 16 - Mapa de relações do modelo de demonstração.

Para este modelo de demonstração foram criadas sete relações entre as diferentes consultas:

- Três relações criadas dizem respeito à relação do atributo “Data” da tabela auxiliar calendário com as datas dos registos das consultas das operações, dos consumos e das colheitas;
- Três relações que associam o atributo “QUADRASnovanomeclatura” da tabela auxiliar de parâmetros físicos aos atributos que dizem respeito às quadras dos registos das operações, dos consumos e da colheita;
- Finalmente, foram associadas as consultas de parâmetros físicos e de previsão de produção através do atributo de identificação de quadras de cada uma das consultas.

Para a elaboração de colunas calculadas relevantes para o conteúdo do modelo é necessário planejar a informação que se pretende transmitir ao utilizador final. Desta forma, o tratamento analítico das consultas contidas no modelo será eficiente e a linguagem informática a introduzir será mais intuitiva. Consequentemente o processo de formulação do cálculo analítico foi segmentado em função dos painéis finais.

Para garantir a funcionalidade do modelo, existe um trabalho extensivo de formulação de colunas calculadas e de medidas na forma da linguagem informática DAX, fundamentais para a correta atividade analítica do modelo. No entanto, a sua referência detalhada nesta dissertação seria injustificada, pelo que de seguida estão somente enunciadas as operações principais realizadas em cada consulta e a sua funcionalidade.



## **Primeiro painel – Produção vs. Previsão ASF**

No primeiro painel, o objetivo principal de demonstração corresponde à comparação entre os valores de previsão de produção e os valores da colheita realizada, na exploração com designação ASF, a diferentes escalas de granularidade respeitantes aos parâmetros físicos.

Este objetivo requer a análise das consultas de colheita, de previsões de produção e dos parâmetros físicos.

Na consulta de parâmetros físicos foi elaborada a medida “Área Total (Ha)” que corresponde à soma das áreas de todas as quadras, possibilitando a filtragem posterior em todos os níveis de granularidade presentes nesta consulta, que neste caso são os níveis de vinha, quadra e variedade.

Na consulta de previsão de produção foi criada a medida “Previsão Total” que consiste na soma dos valores do atributo “2017-2018 PREVISÃO TOTAL”. Esta medida calcula a soma total do peso previsto na colheita, ao nível de granularidade da quadra, permitindo também a sua filtragem em todos os níveis de granularidade existentes na consulta de parâmetros físicos. Foi criada posteriormente a medida “Previsão/Ha” que executa o coeficiente entre a medida “Previsão Total” e a medida “Área Total (Ha)”, igualmente com a possibilidade de ser filtrada em todos os níveis de granularidade.

Na consulta de colheita foi elaborada uma medida de soma dos valores do atributo “PesoLíquido” com um filtro para as colheitas realizadas somente pela exploração com a designação ASF. A esta medida foi dado o nome “Peso Total ASF”. À semelhança da medida “Previsão Total”, esta medida também permite a sua filtragem em todos os níveis de granularidade existentes na consulta de parâmetros físicos. Posteriormente foi criada a medida “Produção/Ha” que corresponde à divisão entre a medida “Peso Total ASF” e a medida “Área Total (Ha)”.

## **Segundo painel – Consumo de água**

No segundo painel deste modelo, o objetivo consiste na passagem de informação ao utilizador final sobre o consumo de água durante três anos distintos, permitindo essa análise a níveis de granularidade distintos.

Para este painel as consultas requeridas consistem nas consultas de consumos, de parâmetros físicos e de calendário.

Neste caso foi necessária a criação de medidas na consulta de consumos, uma vez que já tinha sido previamente calculada a medida que devolve a área total ao nível de granularidade mais reduzido.

Para os objetivos deste painel foram elaboradas duas medidas principais, a medida “Consumo de Água (L)” e a medida “Consumo de Água (mm)”.

A medida “Consumo de Água (L)” consiste na soma dos valores do atributo “QUANT”, da consulta de consumos, filtrando somente os registos associados ao consumo de água.

A medida “Consumo de Água (mm)” diz respeito ao cálculo do consumo de água em milímetros a partir dos valores da medida “Consumo de Água (L)” e dos valores da medida “Área Total (Ha)”, proveniente da consulta dos parâmetros físicos. Esta medida é também criada de forma a ser sensível aos filtros de mudança de granularidade.

### **Terceiro painel – Colheita e supervisores**

No terceiro painel de demonstração é pretendida a exposição dos valores da colheita em função dos supervisores intervenientes nessa operação. Este painel tem como objetivo a comparação de desempenhos de supervisores ao longo da campanha e mostrar alguns indicadores relacionados com a tipologia da colheita e com a marca da exploração responsável pela colheita.

Para a elaboração deste painel é necessária a utilização da consulta de colheita e da consulta de calendário.

As medidas criadas para este painel foram realizadas na consulta de colheita, uma vez que esta possui a totalidade de atributos a abordar.

Foram criadas três medidas principais para a utilização no painel. A medida “Peso Total” que consiste na soma da totalidade do peso líquido de uvas colhidas, a medida “Peso Granel” que soma o peso líquido da colheita realizada pela tipologia a granel e a medida “Peso Embalado Campo” que soma o peso líquido da colheita realizada pela tipologia de embalagem no campo.

Todas as medidas criadas para este modelo partilham a capacidade de serem posteriormente filtradas durante a análise do painel, sem que com isso percam a sua exatidão.

### **Quarto Painel – Registo histórico de consumos**

Tendo em conta a dificuldade de análise de questões relacionadas com o registo de consumos de insumos, o painel de registo histórico de consumos serve o propósito de demonstrar a celeridade e versatilidade de consulta de indicadores relacionados com os consumos que o Power BI® possibilita.

Na criação deste painel foram utilizadas as consultas de consumos, de parâmetros físicos e de calendário.

Depois de algumas operações de uniformização de dados, as principais medidas calculadas foram a “Soma Quantidade” e a “Soma Área”.

A medida “Soma Quantidade” consiste na soma de quantidades dos produtos utilizados, permitindo a filtragem da medida por outros atributos como a data ou a quadra, enquanto a medida

“Soma Área” corresponde à soma de área da exploração, sendo possível a sua filtragem nos vários níveis de granularidade dos dados físicos, por produto utilizado e por data.

#### **Quinto Painel – Registo histórico de operações**

Por fim, o quinto painel permitirá a consulta do histórico de operações realizadas em função de diferentes atributos como o supervisor, a operação realizada, a data ou a granularidade do parâmetro físico desejado. À semelhança do painel anterior, neste é pretendido expor a dinâmica da análise dos dados que é possível realizar no painel através de simples interações com o mesmo.

Na criação deste painel foram utilizadas as consultas de operações, de parâmetros físicos e de calendário.

As medidas principais criadas foram a “Horas Totais Trabalho”, a medida “Dias Necessários”, a medida “Diárias” e a medida “RH’s”. A medida “Horas Totais Trabalho” corresponde à soma de horas utilizadas para executar uma operação, enquanto a medida “Dias Necessários” corresponde à diferença entre o primeiro e último dia de uma determinada operação, contabilizando somente os dias em que se executou essa operação. A medida “Diárias” corresponde ao coeficiente entre a medida “Horas Totais Trabalho” e a duração de um dia de trabalho, em horas. A medida “RH’s” corresponde ao coeficiente entre a medida “Diárias” e a área em que foi realizada a operação, que varia conforme a seleção no painel.

Todas as medidas criadas permitem a sua posterior filtragem durante a análise do painel por parte do utilizador.

#### **4.1.3. Modelo de demonstração**

Como referido, o modelo de demonstração conta com cinco painéis, cada um evidenciando uma forma de tratamento de dados distinta e, consequentemente, com uma finalidade diferente. Todos os painéis apresentados têm como principal característica a interligação entre os visuais, ou seja, qualquer seleção afeta a totalidade do painel, através de funções de filtro. Por outro lado, todos os gráficos visuais expostos também têm a possibilidade de se comportarem como filtros, o que significa que, para além de serem alterados com diferentes seleções, também possuem a capacidade de se comportarem como filtros e alterarem o modelo.

O primeiro painel do modelo (Figura 16) corresponde ao painel de comparação entre a produção e a previsão da mesma, para a campanha de 2017.

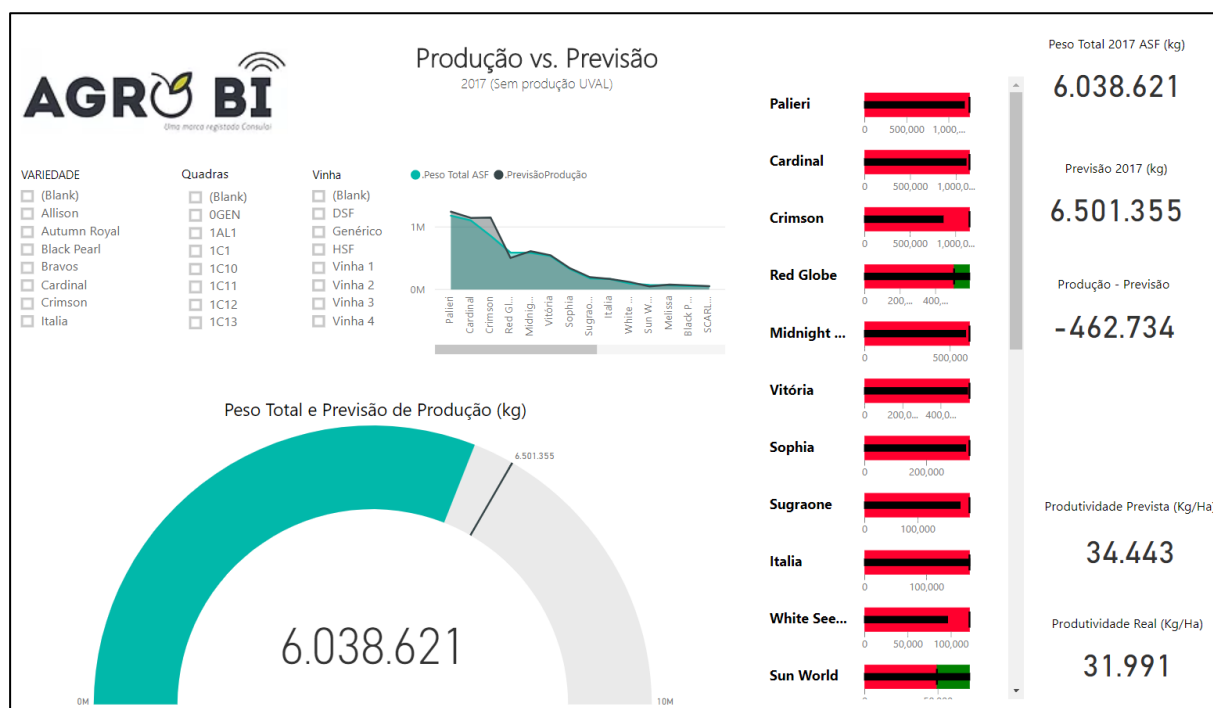


Figura 17 - Painel de comparação entre a previsão de produção e a produção para o ano de 2017.

Na Figura 16 ilustra-se o painel sem nenhuma seleção, o que significa que os valores apresentados nos diferentes gráficos correspondem à totalidade da exploração.

Este painel apresenta três características fundamentais:

- Os filtros, situados no canto superior esquerdo, permitem uma rápida seleção do que pretendemos que seja exibido no painel, ou seja, permite a escolha das diferentes variedades, quadras ou vinhas, sendo que ao ser selecionado um filtro, todo o painel se altera de forma a exibir a informação de acordo com a seleção;
- Os gráficos, situados no centro do painel, exibem informação relacionada com a produção e com a previsão, de uma forma intuitiva. Qualquer um destes é dinâmico conforme os filtros aplicados;
- Do lado direito do painel estão alguns indicadores relevantes, também eles dinâmicos conforme a seleção.

Este painel de comparação visa permitir uma análise crítica do desempenho de cada variedade ou quadra, em função das previsões realizadas no início da campanha.

Na Figura 17 está ilustrado o painel de consumo de água de rega para uma série histórica de registos da empresa.

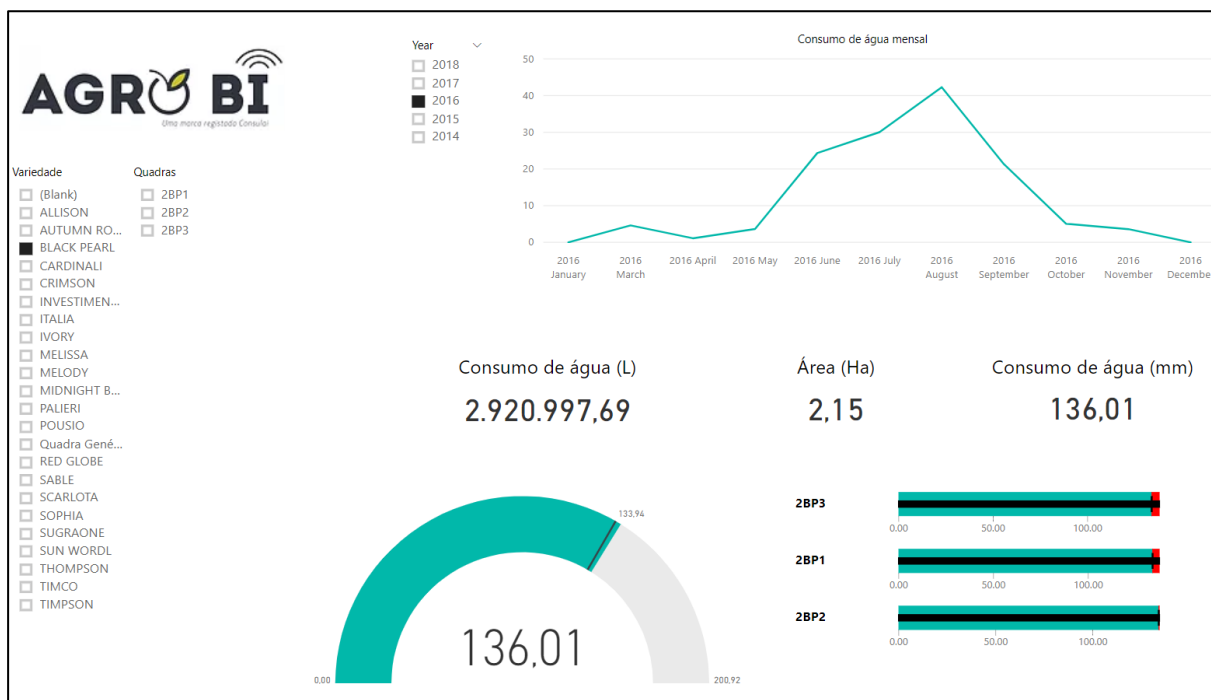


Figura 18 – Painel de consumo de água de rega para diferentes variedades, de finais de 2014 até 2018.

O painel ilustrado na Figura 17 apresenta uma seleção da variedade *Black Pearl* pelo que os dados apresentados correspondem apenas às três quadras que possuem esta variedade.

Este painel tem como principais características três aspetos:

- Os filtros presentes no lado esquerdo do painel, que permitem a seleção da informação que pretendemos visualizar;
- Os gráficos dinâmicos presentes no centro e no canto inferior esquerdo, que permitem ao utilizador interpretar a distribuição da rega ao longo do ano, comparar com a média de consumo de água na totalidade da exploração e comparar o consumo de água entre as diferentes quadras;
- Os indicadores presentes no centro do painel, que revelam indicadores relacionados com o consumo de água em valores absolutos e em valor relativo, de forma a ser possível a comparação entre variedades ou quadras com áreas distintas.

Este painel pretende facilitar a análise do consumo de água de uma forma célere e exata, em função da variedade ou da quadra requerida pelo utilizador.

A Figura 18 apresenta informação relativa aos supervisores de colheita e à colheita realizada.

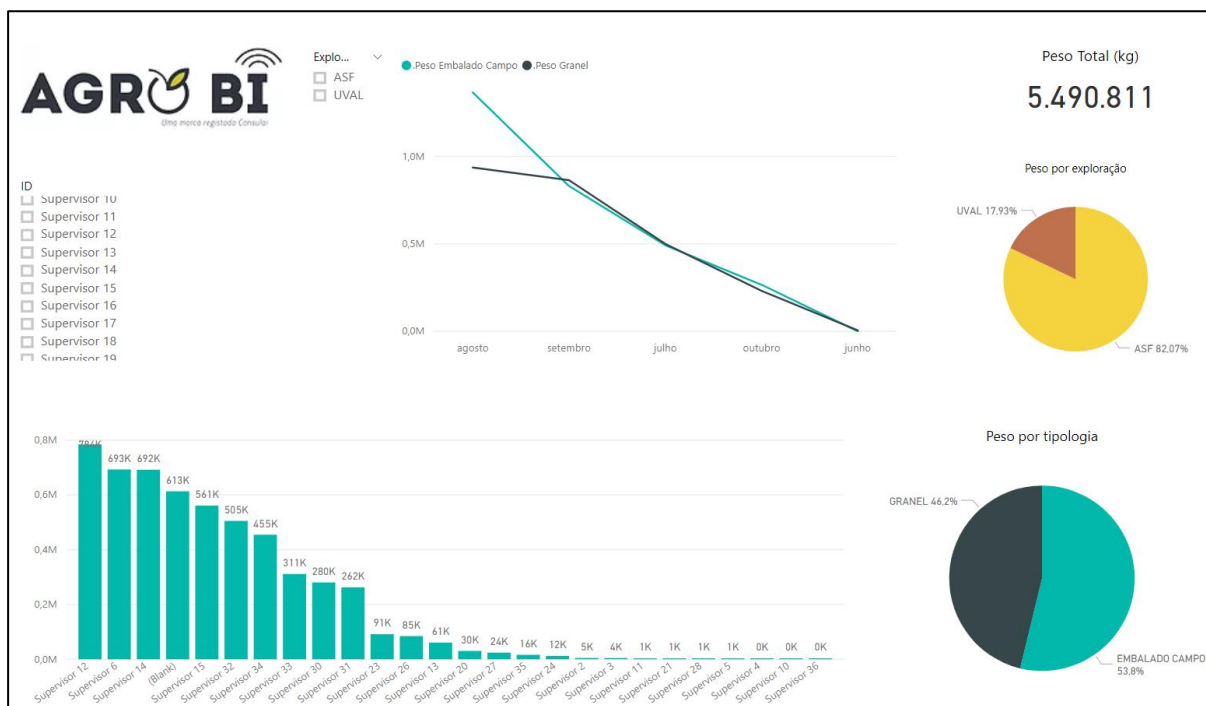


Figura 19 - Painel de análise de supervisores e a sua respetiva colheita durante uma campanha.

Neste caso, o painel ilustra os supervisores presentes na operação colheita e os seus registos em quilogramas de uva colhida. Demonstra também a diferente tipologia de colheita, assim como a marca de exploração responsável pela mesma.

Este painel apresenta dois aspetos principais:

- Os filtros, no lado esquerdo, que permitem a seleção de um ou mais supervisores, de forma a possibilitar a comparação de desempenhos entre eles. Assim como a seleção da marca de exploração;
- Os gráficos, no restante painel, que permitem a comparação entre supervisores, mas também a análise do tipo de colheita feita por cada um e da marca de exploração pela qual foi feita.

Este painel tem como principal utilidade a transmissão de informação sobre o trabalho de colheita realizado pelos supervisores.

O quarto painel (Figura 19) corresponde à análise de registo histórico de consumos de produtos.

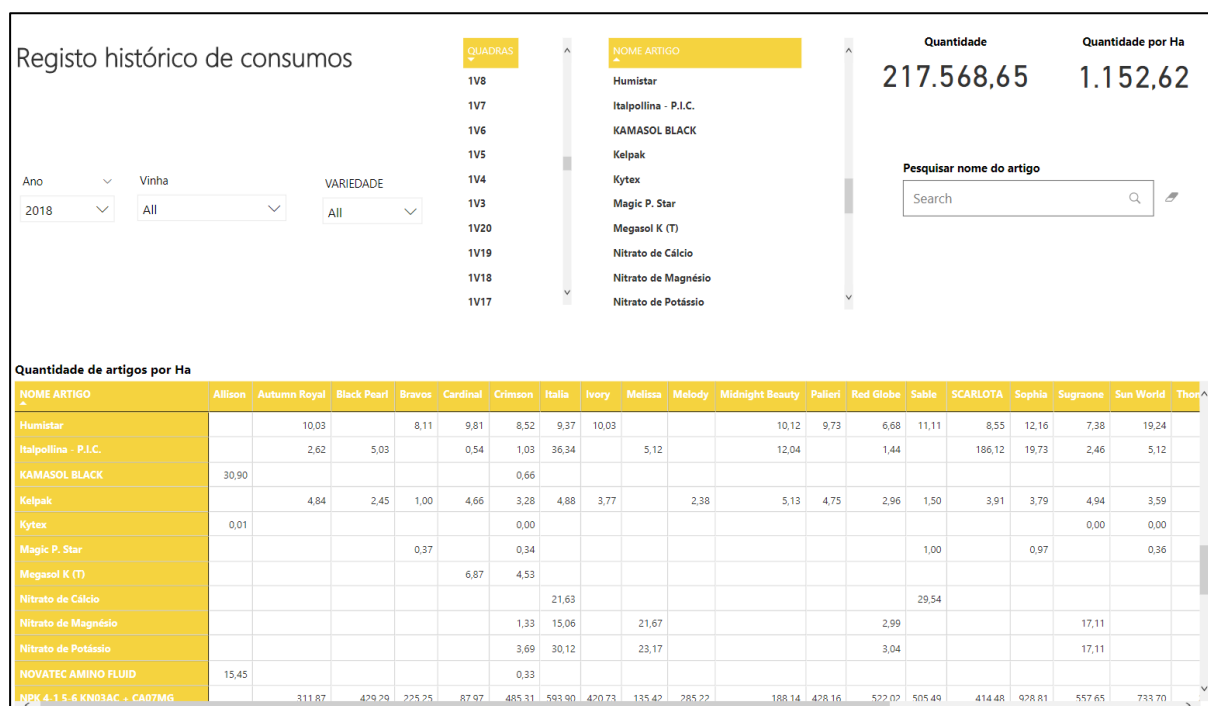


Figura 20 - Paine de registo histórico de consumos

Este painel permite a análise da situação da empresa numa série de registos históricos de consumos, em unidades de volume (litro) por unidade de área (hectare). Tem como características principais:

- Na sua extremidade superior, os filtros, que permitem a seleção dos registos pretendidos, sendo que é possível escrever o nome do produto que se pretende filtrar, uma vez que se tratam de centenas de marcas comerciais de produtos e designações próprias da empresa, o que pode dificultar a sua seleção por listagem;
- Na extremidade inferior, uma tabela dinâmica com os dados de consumos, que relaciona as variedades com os produtos utilizados

O painel da Figura 19 permite assim a comparação de consumos de produtos por unidade de área, entre as diferentes variedades, sendo possível a rápida consulta dos mesmos e comparação entre diferentes anos da série.

Por fim, o painel da Figura 20 corresponde à análise do registo histórico de operações.

Registo Histórico

Ano

2017

Vinha

All

Quadras

All

Variedade

All

Dias Utilizados

309 Dias

Diárias

Operação	1	2	3	4	5	6	7	Gen
ADUBO e FERTILIZANTE	139,39	31,50	79,89	50,93	32,46	74,19	24,76	
ADUBAÇÃO FOLIAR	98,84	20,62	41,38	30,48	26,47	5,86	7,63	
ADUBAÇÃO: Líquida (injectada no solo)	0,44							
ADUBAÇÃO: manual	2,94		0,56		0,25	59,56	11,25	
ADUBAÇÃO: mecanizada	37,18	10,89	37,94	20,45	5,74	8,77	5,89	
AMARRAR CEPA	113,74	33,34	54,34	99,45	31,78	19,91	15,13	
Amarrar enervtos	113,74	33,34	54,34	99,45	31,78	19,91	15,13	
AMARRAR LANÇAMENTOS	197,50		67,00					
Amarrar Lançamentos (+ vezes)	78,37		39,00					
Amarrar Lançamentos (1ª volta)	119,12		28,00					
ATAR VARAS	375,80	109,71	99,88	100,81	80,69	47,40	35,29	
Atar Varas	375,80	109,71	99,88	100,81	80,69	47,40	35,29	
COLOCAR/RETIRAR PLÁSTICOS	4.493,68	597,04	1.558,86	1.721,98	89,06	110,45	108,13	5
Abbrir janelas			6,88	29,88				
Abbrir Panos	13,76	205,07	260,04	16,50				
Atar Panos	529,24	123,45	267,68	285,51	8,63	8,00	24,63	
Atar Redes		2,00	13,31	4,08				
Colocar Lonas			25,73	24,15				
Colocar Panos	900,06	43,83	226,19	309,74	80,38	65,00	30,56	
Colocar Redes	165,56			25,75				
Coser redes	273,81							
Distribuição de plásticos	58,64	1,38	13,97	19,56		0,00	0,88	
Levantar e Baixar saias	15,29	0,15	2,88	2,46	0,06			
Manutenção de Estufas (Panos e Saia)	292,31	26,68	127,20	145,07		5,00		
Manutenção de Estufas (Panos e saias) 2ª Passagem	163,81	10,48	62,57	62,39				
Recolha de Plásticos	15,25							
Tirar Lonas			24,05	49,07				
Tirar Panos	1.654,28	84,69	362,42	647,54		32,45	52,07	
Tirar Redes	83,06							
Volitar panos	313,52	97,07	153,87	77,35				
Volitar Redes	15,06	2,25	12,06	22,94				
CONSERVAÇÃO SISTEMA REGA	22,50	8,60	3,05	73,19	37,01	4,85		2
Rega - Manutenção	22,50	8,60	3,05	73,19	37,01	4,85		
Total	25.361,92	4.030,99	8.521,98	9.111,67	2.156,00	1.597,47	1.007,51	9,5

RH's (Diárias/dia necessárias)

Operação	1	2	3	4	5	6	7	Genérico	Total
ADUBO e FERTILIZANTE	1	0	1	1	1	3	1		2
ADUBAÇÃO FOLIAR	1	0	0	0	1	0	0		2
ADUBAÇÃO: Líquida (injectada no solo)	0								0
ADUBAÇÃO: manual	0		0	0	7	4			3
ADUBAÇÃO: mecanizada	1	1	2	2	1	1			2
AMARRAR CEPA	3	2	2	2	2	2	2	2	4
Amarrar enervtos	3	2	2	2	2	2	2	2	4
AMARRAR LANÇAMENTOS	22	13							19
Amarrar Lançamentos (+ vezes)	16	13							15
Amarrar Lançamentos (1ª volta)	30	14							25
ATAR VARAS	18	16	11	10	6	16	18		21
Atar Varas	18	16	11	10	6	16	18		21
COLOCAR/RETIRAR PLÁSTICOS	29	14	17	24	9	16	15	5	40
Abbrir janelas			3	10					7
Abbrir Panos	14	19	26	17					24
Atar Panos	18	11	15	15	4	8	8		22
Atar Redes		2	3	2					4
Colocar Lonas			6	8					8
Colocar Panos	23	7	19	16	10	33	15	40	24
Colocar Redes	8		13						8
Coser redes	12								12
Distribuição de plásticos	3	1	2	2		0	1		4
Levantar e Baixar saias	2	0	0	0	0				4
Manutenção de Estufas (Panos e Saia)	8	4	5	8		5			12
Manutenção de Estufas (Panos e saias) 2ª Passagem	7	3	4	5				5	10
Recolha de Plásticos	8								3
Tirar Lonas			8	12				1	12
Tirar Panos	45	14	33	54	16	26			52
Tirar Redes	6							3	5
Volitar panos	18	24	15	13					22
Volitar Redes	8	2	12	6					7
CONSERVAÇÃO SISTEMA REGA	2	3	2	4	2	5			4
Rega - Manutenção	2	3	2	4	2	5			4
Total	95	19	34	37	12	14	9	31	195

Figura 21 - Painel de registo histórico de operações

Este painel permite a consulta de registos históricos de dois indicadores importantes para a tomada de decisão na área das operações: as diárias e os RH's necessários. Estes dois indicadores permitem ao utilizador do painel analisar de uma forma célere, a mão-de-obra utilizada anteriormente para realizar operações ou tarefas. Isto permite a execução de uma estimativa de necessidade de trabalhadores a alocar na campanha que estiver a decorrer. Neste painel existem duas características fundamentais:

- Os filtros na zona superior do painel, que permitem a seleção do ano em estudo, mas também, da variedade, da vinha e da quadra;
- As tabelas na zona inferior do painel, que permitem a análise dos dados, mas também a filtragem por operações ou tarefas no resto do painel.

## 4.2. Determinação das necessidades de informação

Uma vez criado o modelo de demonstração, tornou-se necessário reunir com os responsáveis pelas tomadas de decisão. Esta metodologia permite a explicação das capacidades dos modelos de BI e, posteriormente, a interpretação das necessidades reais para uma tomada de decisão o mais eficiente possível, de acordo com os dados disponíveis para a execução da mesma.

No dia 3 de maio de 2018 foi realizada a reunião de demonstração com os responsáveis da empresa, nos escritórios da Herdade do Vale da Rosa, em Ferreira do Alentejo. Na reunião marcaram presença a responsável pela qualidade do produto e o responsável pelas operações. A demonstração do modelo obedeceu a três pontos principais:



- Explicação dos dados utilizados e dos cenários de tomada de decisão criados;
- Demonstração das funcionalidades de cada painel do modelo;
- Identificação das reais necessidades de informação para cada responsável.

No primeiro ponto abordado, realizou-se uma explicação de cenários hipotéticos para cada um dos painéis do modelo de demonstração, em que a solução para a tomada de decisão seria rapidamente encontrada após uma curta consulta dos painéis.

No segundo ponto, foi demonstrada a forma eficiente de obtenção da informação necessária através de um computador e de um *tablet* tendo os participantes da reunião a oportunidade de experimentar o modelo e realizar as suas próprias consultas.

No terceiro ponto, os responsáveis transmitiram as suas críticas ao modelo, tendo apresentado sugestões para novos modelos que colmassem as suas necessidades específicas.

#### **4.2.1 Diferenciação de modelos**

Após a reunião, depois de exposta a potencialidade desta ferramenta, foi decidida a criação de dois modelos ajustados a necessidades específicas de diferentes áreas: (1) Modelo de consumos na fertirrega; (2) Modelo de operações.

Estes modelos têm como objetivo, não só a tomada de decisão em tempo real, mas também a análise histórica do desempenho dos vários departamentos da empresa. Por esta razão a estrutura dos modelos (1) e (2) é constituída por quatro tipos de informação distintos, concretamente, análise de registo histórico, análise dos objetivos ou planeamento inicial, análise da execução e, por fim, análise de comparação entre os objetivos e a execução.

#### **4.2.2. Necessidades técnicas – Consumos na fertirrega**

Após a definição da estrutura do modelo a realizar, verificou-se que as necessidades específicas de cada área de análise divergiam, pelo que foi necessário distinguir o conteúdo de cada um dos quatro painéis a criar. Na área de consumos de fertirrega estas especificidades são:

- Na análise do registo histórico as necessidades principais são: (1) a exposição dos consumos em unidades de volume por unidade de área; (2) a fácil identificação dos produtos no painel; (3) uma análise dinâmica que permita a consulta da informação a vários níveis de granularidade espacial (vinha, quadra, variedade) e em diferentes anos;
- Na análise de objetivos, o pretendido é: (1) a identificação das quantidades de produto a utilizar em função das diferentes datas dos estados fenológicos da videira; (2) uma análise dinâmica que permita a consulta da informação a vários níveis de granularidade espacial (vinha, quadra, variedade);

- Na análise de execução, os objetivos principais são: (1) a organização dos consumos realizados em função das datas de execução dos mesmos; (2) uma análise dinâmica que permita a consulta da informação a vários níveis de granularidade espacial (vinha, quadra, variedade);
- Na análise de comparação entre objetivos e execução é pretendido realizar a diferença entre os dados de execução e os dados do planeamento inicial de forma a ser possível realizar uma análise crítica do que foi executado. Esta diferença assenta nas mesmas premissas de granularidade espacial dos passos anteriores.

#### **4.2.3. Necessidades na análise de operações**

As necessidades na análise de operações distingue-se das demais por tratar de operações realizadas por uma grande variedade de funcionários, o que implica uma elevada variabilidade de eficiência da realização das operações. Por esta razão, e pela natureza dos dados, o pretendido para os painéis diverge das outras áreas da empresa. No entanto, a estrutura dos painéis para este modelo mantém-se, apresentando quatro painéis para análise. Os objetivos principais para este modelo consistem em:

- No primeiro painel é pretendida uma análise do registo histórico dinâmico que pressupõe: (1) uma análise à mão-de-obra por unidade de área, pela totalidade de níveis de granularidade espacial; (2) distinção dos níveis hierárquicos entre operações e tarefas; (3) uma análise dinâmica que permita estimar o número de funcionários a alocar para uma determinada tarefa, em função do número de dias em que se pretende executar a mesma;
- No segundo painel, painel de objetivos, pretende-se: (1) transformar os coeficientes dos objetivos das operações em valores monetários baseados nos registos históricos; (2) expor os custos por operação num formato de continuidade temporal; (3) permitir a rápida seleção de diferentes unidades de granularidade espacial, de forma a permitir uma análise da totalidade da empresa, mas também uma análise específica a uma determinada operação ou a uma determinada unidade espacial;
- No terceiro painel, painel de execução, é pretendida uma análise semelhante ao ponto anterior. Contudo o painel deve ser mais dinâmico e permitir uma análise rápida sobre a integração dos últimos registos que foram inseridos no sistema, de forma a conferir um enfoque maior na situação atual das operações que estão a ser realizadas;
- Por fim, no quarto painel, a análise da diferença entre objetivos e execução tem como objetivos: (1) apresentar todos os resultados na forma de custos por unidade de área; (2) apresentar o orçamento disponível para operações e tarefas em função da granularidade espacial (vinha, variedade e quadra); (3) existir uma visualização que organize o orçamento em função do calendário, de forma a permitir a fácil análise da situação atual e da margem de fundos que existe ao longo da campanha.

#### **4.4. Modelação final**

Uma vez definidas as necessidades de cada área em estudo da empresa, foi necessária a criação de dois modelos distintos, com procedimentos distintos. Deste modo, a modelação também foi realizada separadamente, em concreto, o modelo de consumos e o modelo de operações.

##### **4.4.1. Modelo de consumos**

O modelo de consumos tem como principal objetivo a exposição de indicadores que auxiliem, em tempo real, o processo de tomada de decisão no consumo de produtos na fertirrega.

Para a elaboração deste modelo foram utilizados os dados de registo de consumos (Figura 4), de planeamento de consumos (Figuras 5 e 6), de estados fenológicos (Figura 11) e de dados físicos (Figura 12).

Para a automatização da utilização de dados, foram criadas pastas nas quais foram introduzidos os registos no seu formato original. Posteriormente, com recurso ao editor de consultas e à linguagem informática M, foi programada a forma como as tabelas incluídas em cada pasta devem ser editadas, de forma a adquirirem o formato da Tabela 2. Este procedimento permite que no futuro, aquando da existência de mais registos, seja somente necessário transferir os ficheiros originais para a pasta fonte do modelo.

De forma a beneficiar as capacidades analíticas do sistema foram criadas três tabelas auxiliares:

- Uma tabela auxiliar calendário, como ilustrada na Tabela 3, que terá como objetivo o enquadramento temporal dos diferentes registos;
- Uma tabela auxiliar de designação de produtos, que serve a finalidade de tradução e correspondência entre as designações de produtos presentes na tabela de objetivos e as designações presentes na tabela de registo de consumos (Figura 21);
- Uma tabela auxiliar de dados físicos por variedade, de forma a facilitar o trabalho analítico em diferentes níveis de granularidade (Figura 21).

DesignaçãoObjectivos	DesignaçãoConsumos
Ácido Humico	Humistar
sulfato de magnésio	sulfato de magnesio
SuperGat Ácido	Adubo 5,6-8-2+6,4+0,8 SA
Sulfato de Potássio	Sulfato de Potássio
Sulfato de Amónio	Sulfato de Amónio
Quelato de Ferro	Ferronova Express
Nitrato de Cálcio	Nitrato de Cálcio
MAP	Fosfato de Monoamonio - Krista map
Adubo 1-1-9	Adubo 1-1-9

VARIETADE	Área (Ha)
Allison	0,9995
Autumn Royal	1,96
Black Pearl	2,1477
Bravos	0,61642
Cardinal	25,475
Crimson	44,156
Italia	4,98
Ivory	1,9667

Figura 22 - Excerto das tabelas auxiliares de designação de produtos (esquerda) e de dados físicos por variedade (direita).

Depois da edição das consultas e criação das tabelas auxiliares foi elaborado o mapa de relações deste modelo (Figura 22). Nesta fase foram criadas oito relações entre as consultas, respetivamente:

- Duas relações da tabela auxiliar calendário para as consultas de consumos e objetivos, respetivamente;
- Duas relações da tabela auxiliar de designação de produtos para as consultas de objetivos de consumos e registo de consumos, respetivamente;
- Duas relações da consulta de dados físicos para a tabela auxiliar de dados físicos (com mais um nível de granularidade) e para a consulta de registo de consumos.
- Uma relação da tabela auxiliar de dados físicos para a consulta de objetivos de consumos;
- Uma relação da consulta de estados fenológicos para a consulta de objetivos de consumos.

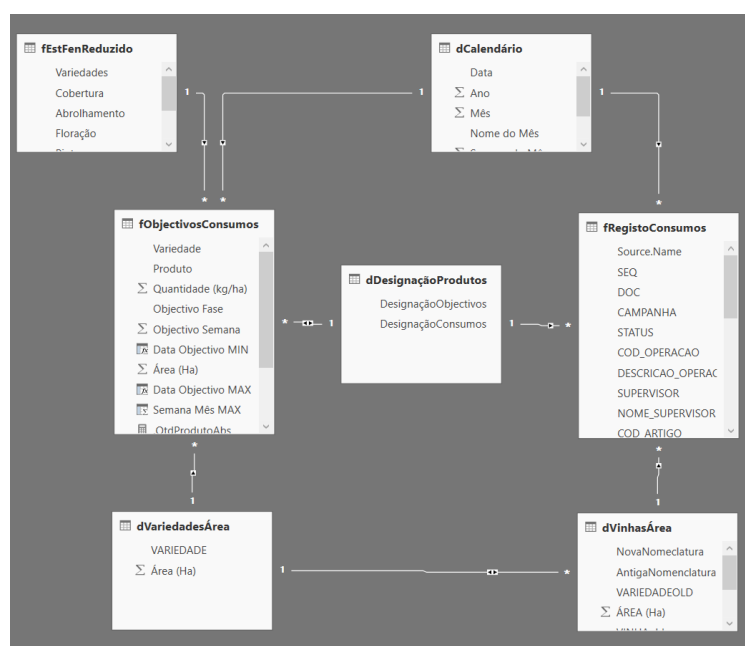


Figura 23 - Mapa de relações do modelo de consumos.

Para a criação deste modelo foram elaborados quatro painéis distintos, que seguem a metodologia proposta na determinação de necessidades com o responsável da empresa pela área de consumos, ou seja, (1) painel de registo histórico, (2) painel de objetivos, (3) painel de execução e (4) painel de comparação entre objetivos e execução.

Nesta fase é necessária a utilização da linguagem informática DAX para a criação de medidas analíticas e colunas calculadas, de forma a permitir a correta exposição de informação nos diferentes painéis, de acordo com o que foi requerido na determinação de necessidades.

### **Primeiro painel – Painel de registo histórico**

Neste painel é pretendida a consulta rápida de consumos, à semelhança do que tinha sido elaborado no painel de registo histórico de consumos no modelo de demonstração, pelo que foram importadas as medidas realizadas no mesmo e utilizadas na criação deste modelo final. Em concreto as medidas criadas foram:

- A medida “Soma Quantidade” que consiste na soma de quantidades dos produtos utilizados, permitindo a filtragem da medida por outros atributos como a data ou a quadra;
- A medida “Soma Área” que corresponde à soma de área da exploração, sendo possível a sua filtragem nos vários níveis de granularidade dos dados físicos, por produto utilizado e por data;
- A medida “Quantidade Produto/ha” que corresponde ao coeficiente entre as medidas “Soma Quantidade e “Quantidade Produto/ha”, partilhando as mesmas possibilidades de filtragem das duas medidas anteriores.

### **Segundo painel – Planeamento inicial de consumos**

Neste segundo painel é pretendida a integração das tabelas de planeamento de consumos, de uma forma dinâmica, com os registos físicos, de forma a ser possível consultar o que deve ser aplicado, quando deve ser aplicado, onde deve ser aplicado e que produto deve ser aplicado.

Para esta finalidade as principais operações realizadas foram a criação de uma coluna calculada e de uma medida, respetivamente:

- A coluna “Data Objetivo min”, que consiste na integração das datas das fases fenológicas das variedades (Figura 11), juntamente com o planeamento de consumos (Figura 5 e 6). Tendo em conta que o planeamento de consumos está elaborado em função do início de cada fase fenológica e intercalado por acréscimos semanais, esta coluna fornece a data do início da semana de transição de estado fenológico;

- A medida “Quantidade de Produto”, que corresponde à multiplicação dos consumos planeados em unidades de volume por unidade de área, pela área da seleção realizada. Esta medida permite ser distribuída temporalmente resultando assim num valor de consumo de produtos por unidade de tempo, neste caso o mês ou a semana, ou ambos.

### **Terceiro painel – Execução de consumos**

No terceiro painel é pretendida a exposição da realidade atual da empresa respeitando o consumo de produtos, o que se traduz num painel de registo histórico do presente ano, apresentado com indicadores relevantes para uma tomada de decisão atual.

Para essa finalidade não foi necessária a criação de novas medidas ou colunas calculadas, mas sim uma exposição diferente do primeiro painel. As medidas utilizadas foram a “Quantidade Produto/ha” e a “Soma Quantidade”.

### **Quarto painel – Comparação de planeamento e execução de consumos**

No quarto e último painel deste modelo, o pretendido é permitir a análise em tempo real da diferença entre o planeamento e execução de consumos de produtos, em função dos diferentes dados físicos e no decorrer da campanha.

Para essa finalidade foi necessária a criação de quatro novas medidas, respetivamente:

- A medida “YTD Execução”. A sigla YTD (*Year To Date*) corresponde a uma soma cumulativa anual. Neste caso essa soma cumulativa corresponde aos produtos consumidos;
- A medida “YTD Planeamento”, consiste numa soma cumulativa anual dos consumos planeados;
- A medida “%Dif Planeamento & Execução”, consiste na diferença entre as duas medidas anteriores, na forma de percentagem. Sendo que valores superiores a 100% significam que a execução foi superior ao planeado e valores inferiores a 100% significam que a execução foi inferior ao planeado.

#### **4.4.2. Modelo de operações**

A modelação do modelo de operações almeja auxiliar a análise das operações realizadas, assim como auxiliar a tomada de decisão em tempo real na área das operações da empresa.

Para a elaboração do modelo de operações foram utilizados dados correspondentes às tabelas de registo de operações (Figura 7), de objetivos de operações (Figura 8) e de dados físicos (Figura 12).

Para a automatização da utilização de dados, foram criadas pastas nas quais foram introduzidos os registos no seu formato original. Posteriormente, com recurso ao editor de consultas e à linguagem informática M, foi programada a forma como as tabelas incluídas em cada pasta devem ser editadas, de forma a adquirirem o formato da Tabela 2. Este procedimento permite que no futuro, aquando da existência de mais registos, seja somente necessário transferir os ficheiros originais para a pasta fonte do modelo.

De forma a permitir o correto funcionamento do programa e permitir uma clara exposição de resultados, foi necessário recorrer à criação de três tabelas auxiliares no editor de consultas, concretamente:

- A tabela “Número de trabalhadores” que consiste numa lista de contagem de intervalo de 5 unidades até um máximo de 200 unidades (Figura 23). Esta tabela cumprirá a função de auxiliar uma visualização num painel de análise ao registo histórico das operações;
- A tabela “Tarefa” que corresponde à atribuição de códigos de operação às diferentes tarefas (Figura 24). Esta tabela servirá para auxiliar as mudanças de granularidade feitas no painel pelos utilizadores;
- A tabela “Calendário” como ilustrada na Tabela 3. Que servirá o propósito de enquadramento temporal dos diferentes registos.

NúmeroTrabalhadores
0
5
10
15
20
25
30
35
40
45

Figura 24- Excerto da tabela auxiliar "Número de trabalhadores".

CODIGO FAMILIA	CODIGO TAREFA	TAREFA	Operação
C.1	1	Marcação da Poda Seca	C.1 - MARCAÇÃO
C.4	2	Atar Varas	C.4 - ATAR VARAS
C.2	3	Poda Seca	C.2 - PODA SECA
C.3	4	Destroçar Varas	C.3 - DESTROÇAR VARAS
G.1	8	Colocar Panos	G.1 - COLOCAR/RETIRAR PLASTICOS
G.1	9	Colocar Redes	G.1 - COLOCAR/RETIRAR PLASTICOS
G.1	10	Tirar Panos	G.1 - COLOCAR/RETIRAR PLASTICOS
G.1	11	Tirar Redes	G.1 - COLOCAR/RETIRAR PLASTICOS
B.3	13	Amarrar enxertos	B.3 - AMARRAR CEPA
E.1	14	Tirar Folhas Basilares	E.1 - TIRAR FOLHAS
G.3	15	Manutenção - arame	G.3 - CONSERVAÇÃO VINHA
E.2	16	Seleção de Cachos	E.2 - SELECÇÃO DE CACHOS
D.2	17	Incisão no tronco	D.2 - INCISÃO

Figura 25 - Excerto da tabela auxiliar "Tarefa".

Posteriormente à edição das consultas e criação das tabelas auxiliares foi elaborado o mapa de relações deste modelo (Figura 25). Nesta fase foram criadas quatro relações entre as consultas, concretamente:

- Uma relação entre a tabela auxiliar “Calendário” e a consulta de registo de operações;
- Uma relação entre a tabela auxiliar “Tarefa” e a consulta de registo de operações;
- Uma relação entre a consulta de dados físicos e a consulta de registo de operações;
- Uma relação entre a consulta de objetivos de operações e a tabela auxiliar “Tarefa”.

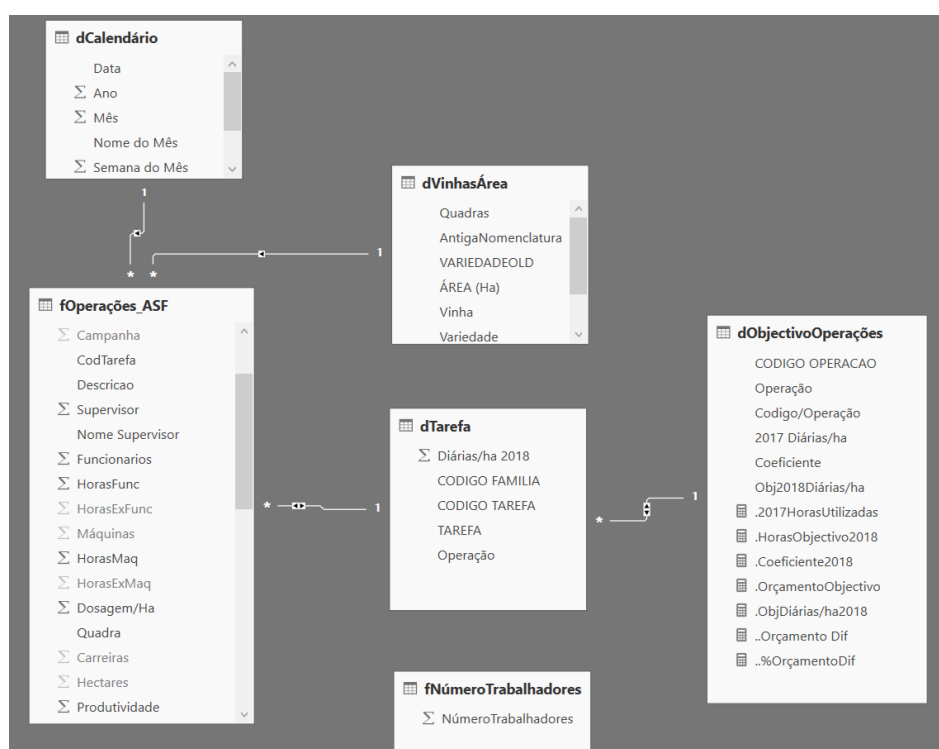


Figura 26 - Mapa de relações do modelo de operações.



Para a elaboração deste modelo foram criados cinco painéis distintos, que seguem a metodologia proposta na determinação de necessidades com o responsável da empresa pela área de operações, ou seja, (1) painel de registo histórico, (2) painel de registo histórico dinâmico, (3) painel de objetivos, (4) painel de execução e (5) painel de comparação entre objetivos e execução.

Nesta fase é necessária a utilização da linguagem informática DAX para a criação de medidas analíticas e colunas calculadas, de forma a permitir a correta exposição de informação nos diferentes painéis, de acordo com o que foi requerido na determinação de necessidades.

### **Primeiro painel – Painel de registo histórico**

No painel de registo histórico é pretendida uma consulta simples e célere de alguns indicadores relacionados com as operações. Dado que só será necessária a exibição simples dos mesmos, foram criadas as seguintes medidas:

- “Horas Totais Trabalho”, que consiste numa soma de horas de trabalho de todos os registos, permitindo a sua filtragem por atributos como o ano, variedade, vinha, quadra, operação, tarefa e operador;
- “Dias Necessários”, que consiste numa soma de dias distintos em que as operações ou tarefas foram realizadas;
- “Diárias”, que é constituída pelo coeficiente entre a medida “Horas Totais de Trabalho” e a duração de um dia de trabalho em horas. Esta medida também possui a capacidade de ser filtrada pelos mesmos atributos que a medida “Horas Totais Necessárias”;
- “RH's”, que representa coeficiente da medida “Diárias” pela área que se encontra selecionada no painel. Em caso de não existir seleção, a medida assume a totalidade da área da exploração por defeito.

### **Segundo painel – Painel de registo histórico dinâmico**

No segundo painel deste modelo, é pretendida uma análise dinâmica do registo histórico de forma a ser possível criar pequenos cenários que auxiliem a tomada de decisão aquando a alocação de equipas para determinadas tarefas ou operações. Para tal foram criadas as seguintes medidas:

- “Duração da Operação” que corresponde ao número de dias que foram utilizados para a realização de uma operação;
- “Inicio e Fim de Operação” que consiste na diferença de datas entre o fim e o início de uma operação, em dias. Esta medida distingue-se da medida “Duração da Operação” por contabilizar a totalidade de dias entre as duas datas e não somente os dias em que a operação se realizou;

- “Horas para exposição” que corresponde a uma medida de cariz complexo, ou seja, calcula mais do que um valor. A sua função é de exibir o valor da medida “RH’s” juntamente com a operação selecionada, na forma de texto. Se não existir seleção, esta medida assume por defeito os dados de todos os registos, produzindo o texto “para todas as operações”.

### **Terceiro painel – Objetivos de operações**

No terceiro painel é pretendida a exibição dinâmica dos objetivos das operações e tarefas para a corrente campanha, tendo em consideração os coeficientes de objetivos que foram cedidos. Para tal foram desenvolvidas as seguintes medidas:

- “Horas Utilizadas 2017” que corresponde à soma de horas utilizadas por tarefa no ano de 2017;
- “Horas Objetivo 2018”, que consiste na multiplicação do coeficiente de eficiência pela medida “Horas Utilizadas 2017”;
- “Orçamento Objetivo 2018” que corresponde à multiplicação da medida “Horas Objetivo 2018” pelas remunerações horárias dos funcionários.

### **Quarto painel – Execução de operações**

O painel de execução das operações pretende permitir a análise do custo das operações na corrente campanha na exploração de uma forma dinâmica, ou seja, que relacione os dados físicos, permitindo a identificação e comparação rápida dos diferentes custos operacionais. Para essa finalidade foram criadas quatro medidas principais:

- A medida “Custo Corrente” que executa a multiplicação entre a medida “Horas Totais de Trabalho” e as remunerações horárias dos funcionários;
- A medida “Custo Corrente/ha” que executa o coeficiente entre a medida “Custo Corrente” e a medida “Soma Áreas”;
- A medida “Custo Corrente Ha Fixo” que consiste na medida “Custo Corrente/ha”, mas, no entanto, é insensível aos filtros do painel, apresentando de uma forma constante o custo operacional por hectare de toda a exploração;
- A medida “Custo Max Gauge” que corresponde ao valor máximo anual dos últimos três anos do registo da medida “Custo Corrente/ha”.

### **Quinto painel – Comparação entre objetivos e execução de operações**

O quinto e último painel pretende permitir a análise da situação das operações da empresa, comparando os custos operacionais executados e os custos operacionais planeados no início da

campanha, em tempo real e ao longo do ano. De forma a atingir esta finalidade foi necessária a criação de duas medidas:

- A medida “Orçamento Dif”, que calcula a diferença entre as medidas “Orçamento Objetivo 2018” e “Custo Corrente”, resultando assim num saldo em que, em caso de valores negativos, o orçamento foi ultrapassado, e em casos de valores positivos, a execução foi realizada dentro das margens orçamentais planeadas. Esta medida tem a possibilidade de ser filtrada, permitindo a análise em todos os níveis de granularidade de dados físicos e dos dados operacionais.

## 4.5. Modelos finais

Uma vez realizadas todas as operações de edição de consultas e de operações analíticas foram elaborados os painéis para cada modelo. Desta fase resultaram os modelos finais disponíveis para integração na empresa. Como referido anteriormente, foram concebidos dois tipos de modelos distintos, (1) modelo de consumos e (2) modelo de operações.

### 4.5.1. Modelo de consumos

O primeiro painel do modelo de consumos corresponde ao registo histórico de consumos (Figura 26). Tendo em consideração os resultados da reunião de determinação de necessidades, foi decidido que este painel original do modelo de demonstração integraria o modelo de consumos final.

Registo histórico de consumos

Ano: 2018 Vinha: All VARIEDADE: All

QUADRIAS: HSF\_CR3, HSF\_CR2, HSF\_CR1, DSF\_CR3, DSF\_CR2, DSF\_CR1, STS1, SIV1, SEX1, SCR9

NOME ARTIGO: Acido Fosfórico, Acido Nitrico, Active Erger, Adubo 14-7-17, AEGIS SYM GEL ARGILA, AGRO ENHANCER MG, ALGAMAN B, Algaren Conc. L, Amino-quelant CA

Quantidade: 217.568,65 Quantidade por Ha: 1.152,62

Pesquisar nome do artigo: Search

Quantidade de artigos por Ha

NOME ARTIGO	Allison	Autumn Royal	Black Pearl	Bravos	Cardinal	Crimson	Italia	Ivory	Melissa	Melody	Midnight Beauty	Pallari	Red Globe	Sable	SCARLOTA	Sophia	Sugraone	Sun World
Acido Fosfórico	2,38				0,04	0,27			0,04								1,55	0,14
Acido Nitrico													1,29					
Active Erger									93,75								26,22	
Adubo 14-7-17	262,63	411,49		191,17		51,20	355,88		362,67				284,26		303,65		33,95	
AEGIS SYM GEL ARGILA							0,26											
AGRO ENHANCER MG					1,30								0,95					
ALGAMAN B																		
Algaren Conc. L											0,50							
Amino-quelant CA																	0,65	
ARBOKOL						0,02		5,09										
Aton az		1,60		0,30	1,56	0,42	1,65	1,16				1,35	1,42	1,04		1,32	1,56	1,10
AVANT NATURE	1,80					0,04												

Figura 27 - Painel de registo histórico de consumos.

Este painel permite a análise da situação da empresa numa série de registos históricos de consumos, unidades de volume (litro) por unidade de área (hectare). Tem como características principais:

- Na sua extremidade superior, os filtros, que permitem a seleção dos registos pretendidos, sendo que é possível escrever o nome do produto que se pretende filtrar, uma vez que se tratam de centenas de marcas comerciais de produtos e designações próprias da empresa, o que pode dificultar a sua seleção por listagem;
- Na extremidade inferior, uma tabela dinâmica com os dados de consumos, que relaciona as variedades com os produtos utilizados.

O painel da Figura 26 permite assim a comparação de consumos de produtos por unidade de área, entre as diferentes variedades, sendo possível a rápida consulta dos mesmos e comparação entre diferentes anos da série.

O segundo painel consiste no plano inicial de consumos para 2017 (Figura 27).

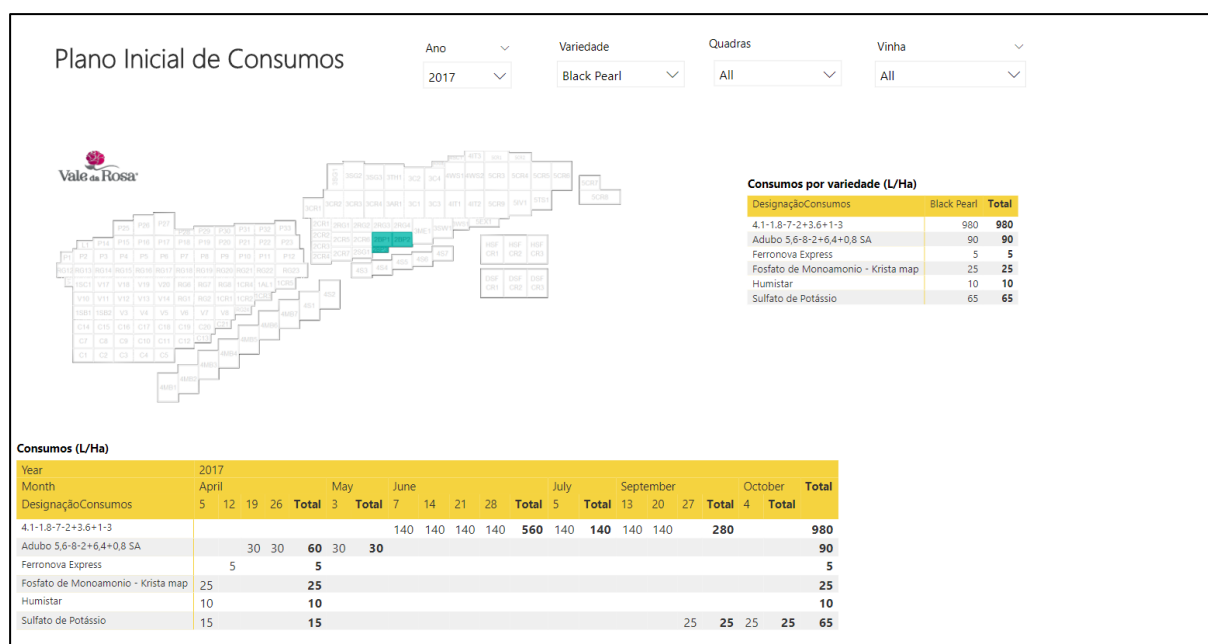


Figura 28 - Painel de plano inicial de consumos para 2017.

Este painel permite ao utilizador consultar os consumos de produtos planeados para a campanha. Na Figura 27 em particular encontra-se selecionada a variedade *Black Pearl* de forma a demonstrar a dinâmica dos quadros. Estes alteram o seu conteúdo conforme os filtros selecionados, sendo assim possível uma consulta rápida dos produtos que se planeiam consumir e da distribuição de consumos durante o ano. Em detalhe, este painel é composto por:



À semelhança do painel de objetivos, o painel de execução de consumos apresenta dois quadros com valores de consumos. No entanto, neste caso, os consumos expostos são realmente os executados, em concreto, os consumos executados na vinha 1 ao longo da campanha. Em concreto, este painel tem como principais componentes:

- Na extremidade esquerda, os filtros de seleção, que permitem o ajuste do painel às necessidades do utilizador;
- No centro, dois quadros que exibem informação sobre os consumos executados. O quadro inferior exibe informação em unidades de volume (litros) por unidade de área (hectare), enquanto o quadro superior exibe informação em valores absolutos, em litros.

O painel de diferença entre execução e plano inicial de consumos encontra-se representado na Figura 29.

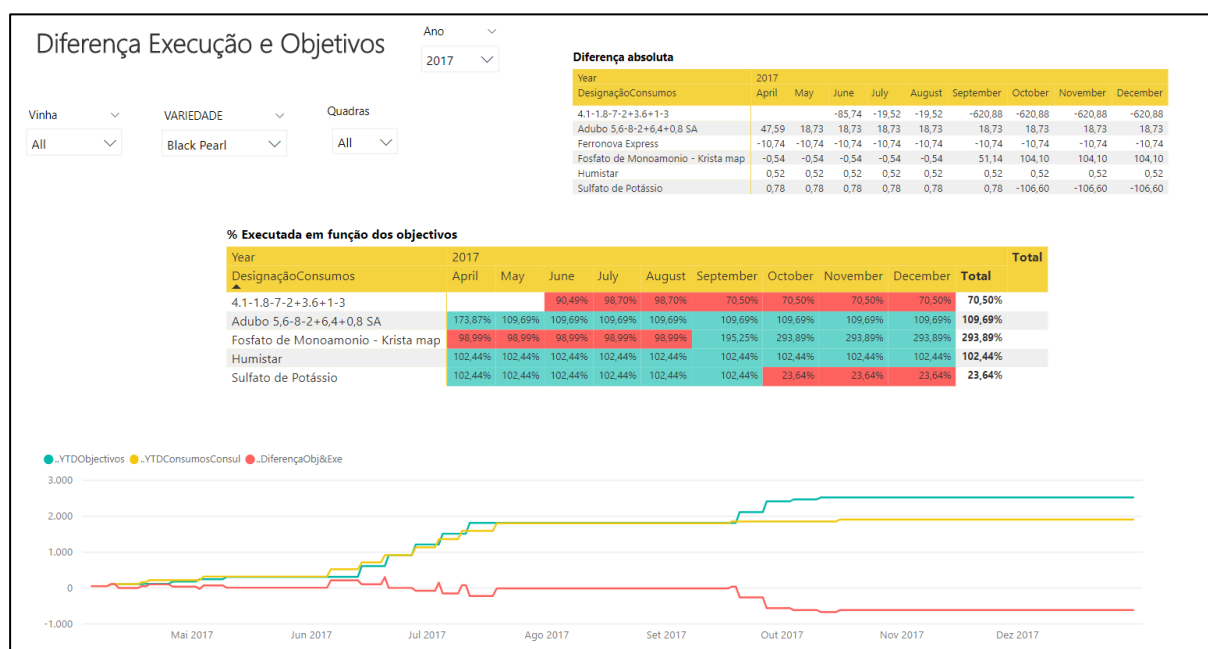


Figura 30 - Painel de diferença entre execução e plano inicial de consumos.

Este quarto e último painel do modelo de consumos pretende comparar as decisões realizadas na execução de consumos juntamente com o plano realizado no início da campanha. Especificamente na Figura 29 encontra-se selecionada a variedade *Black Pearl* mais uma vez. Desta forma é possível interpretar que todo o painel se encontra ajustado ao caso específico da variedade selecionada. Este painel permite assim a interpretação do desempenho da execução, comparada com o planeamento, na forma absoluta, em percentagem e graficamente. Todas estas análises estão expostas no espetro temporal da campanha realizada. Este painel conta assim com quatro pontos principais:

- No canto superior direito estão presentes os filtros de seleção que permitem ajustar a informação exposta pelo painel em função das necessidades do utilizador;

- Na extremidade superior, o quadro de valores absolutos (litros) da diferença entre planeamento e execução, em que os valores negativos se traduzem numa execução de consumo que excedeu o plano inicial;
- No centro do painel encontra-se o quadro de diferença entre o planeamento e execução em percentagem, em que os valores superiores a 100% significam que a execução excedeu o consumo planeado no início da campanha;
- Finalmente, na extremidade inferior está presente um gráfico que mostra as curvas de execução de consumos, planeamento de consumos e a curva de diferença entre os dois. Este gráfico tem como objetivo a exposição gráfica do período em que os desvios entre as duas curvas decorreram.

#### 4.5.2. Modelo de operações

Neste capítulo encontram-se ilustrados os painéis resultantes de todo o processo de modelação dos dados relativos às operações na Herdade Vale da Rosa.

O primeiro painel (Figura 30) diz respeito à análise do registo histórico das operações executadas no período compreendido entre 2014 a 2018 (campanhas de 2015 a 2018).

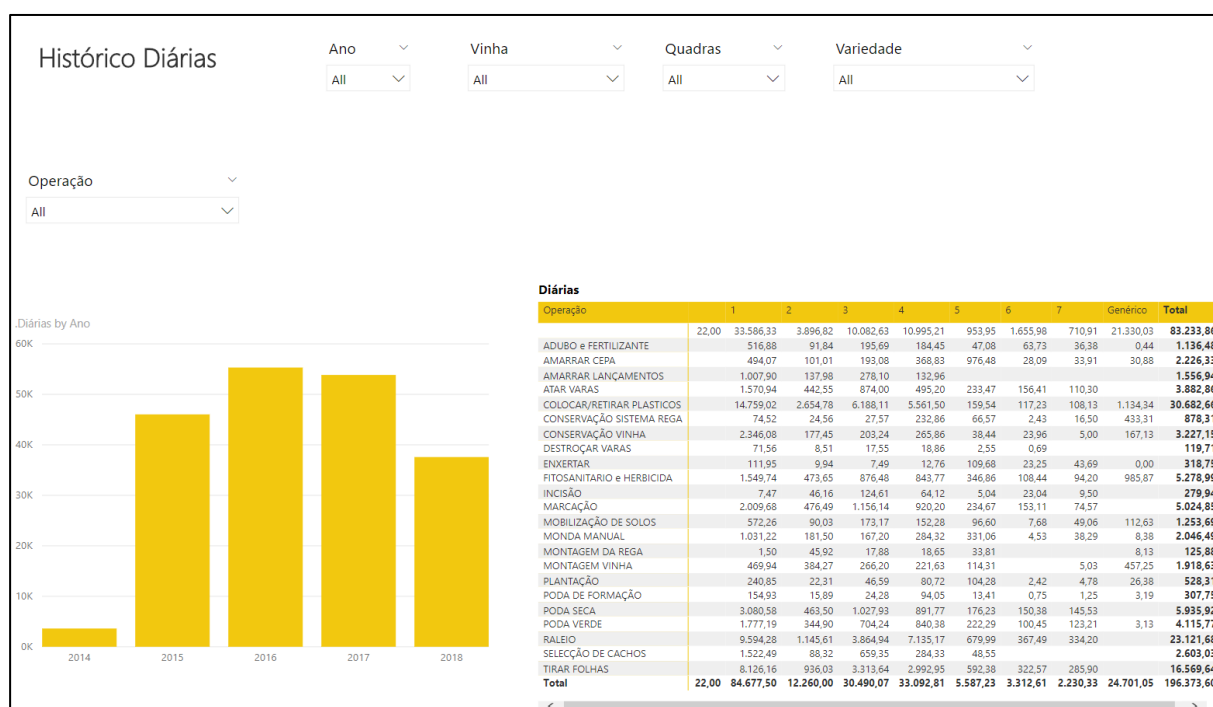


Figura 31 - Painel de registo histórico de operações

Neste painel não se encontra selecionado nenhum ano em concreto, pelo que todo o conteúdo exposto é respeitante a toda a série histórica. O conteúdo apresentado possui três principais elementos:

- Na extremidade superior, os filtros que permitem uma seleção que melhor se adequa às necessidades do utilizador;
- Na extremidade esquerda, um gráfico que permite a comparação da medida “Diárias” por hectare nos diferentes anos, nas diferentes seleções que podem ser realizadas;
- Na extremidade direita, uma tabela que apresenta a medida “Diárias” em função das operações (e tarefas) e da vinha (ou uma outra granularidade física pretendida).

Ao permitir a seleção de vários níveis de granularidade ao nível da operação e dos dados físicos, este painel tem como objetivo exibir uma análise operacional detalhada ao utilizador e a comparação da mesma nos diferentes anos.

Na Figura 31 apresenta-se um outro painel com incidência no registo histórico, denominado painel de histórico dinâmico.

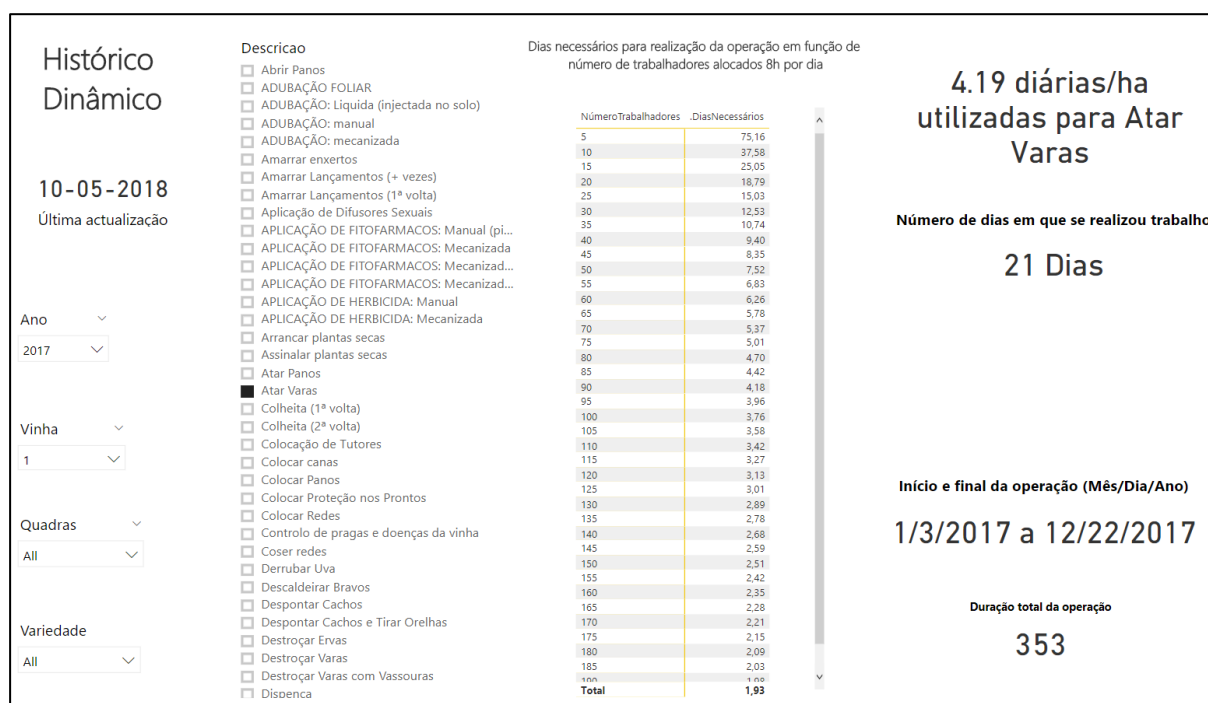


Figura 32 - Painel de registo histórico dinâmico de operações.

Este segundo painel difere do primeiro ao nível da exposição da informação, permitindo uma interpretação distinta requerida pelo responsável pelas operações da empresa. Para esse fim, o painel dispõe de quatro principais características, designadamente:

- Os filtros na extremidade esquerda, que permitem uma seleção a vários níveis de granularidade dos dados expostos;
- A listagem das tarefas, que também pertence à categoria de filtros, mas neste caso desempenha uma função de filtro principal do painel;
- A listagem de número de trabalhadores alocados e o número de dias necessários para completar a tarefa selecionada. Esta listagem analisa a série histórica de dados e devolve uma



estimativa de dias necessários para completar a tarefa selecionada, em função do número de trabalhadores que realizam a tarefa;

- Os valores na extremidade esquerda, que auxiliam o utilizador na interpretação dos resultados.

O painel de registo histórico tem como objetivo principal o auxílio na análise de necessidades de recursos humanos para diferentes tarefas, o que por sua vez irá auxiliar a tomada de decisão de alocação de mão-de-obra nas diferentes tarefas, mas também em diferentes parcelas ou variedades.

O terceiro painel deste modelo (Figura 32) corresponde aos objetivos anuais das operações, baseados no ano anterior.

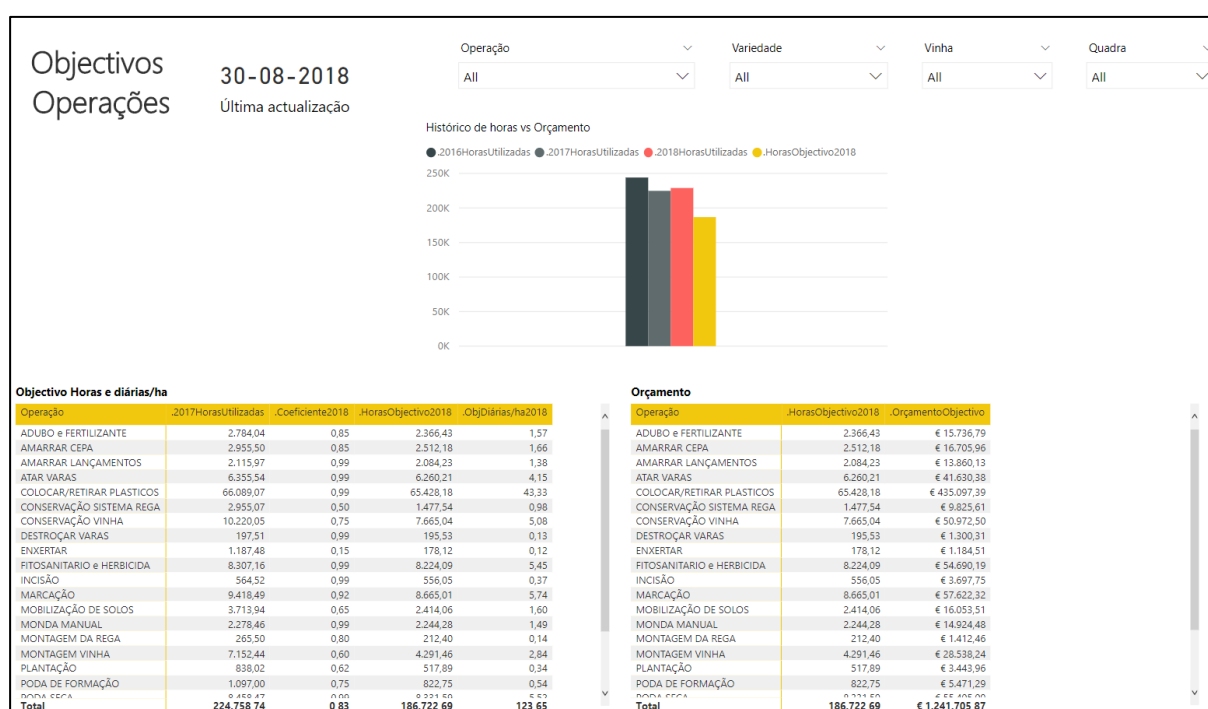


Figura 33 - Painel de objetivos das operações.

O painel de objetivos das operações tem como objetivo fundamental permitir uma análise crítica dos objetivos propostos, na forma de coeficientes, atribuindo aos objetivos valores na forma de “Diárias por hectare” e na forma de custo em euros.

Este painel expõe assim dois quadros e um gráfico, que revelam um orçamento para a campanha corrente, especificamente:

- Na extremidade superior do painel estão presentes os filtros de seleção de forma a permitir uma consulta ajustada às necessidades do utilizador;
- No centro encontra-se um gráfico que ilustra informações relativas à utilização de horas nos anos de 2016, 2017 e 2018, juntamente com as horas projetadas para 2018, respetivamente;

- Na extremidade inferior, existem duas tabelas que permitem a consulta dos coeficientes de objetivos para as diferentes operações e tarefas, assim como as horas, as diárias por hectare e o custo de remuneração da mão-de-obra.

O painel de objetivos tem por finalidade o auxílio da programação das operações a realizar na corrente campanha, de forma congruente com os objetivos propostos.

A Figura 33 ilustra o painel de execução de operações. Neste painel são tratados os dados de operações com possibilidade de atualização em tempo real.

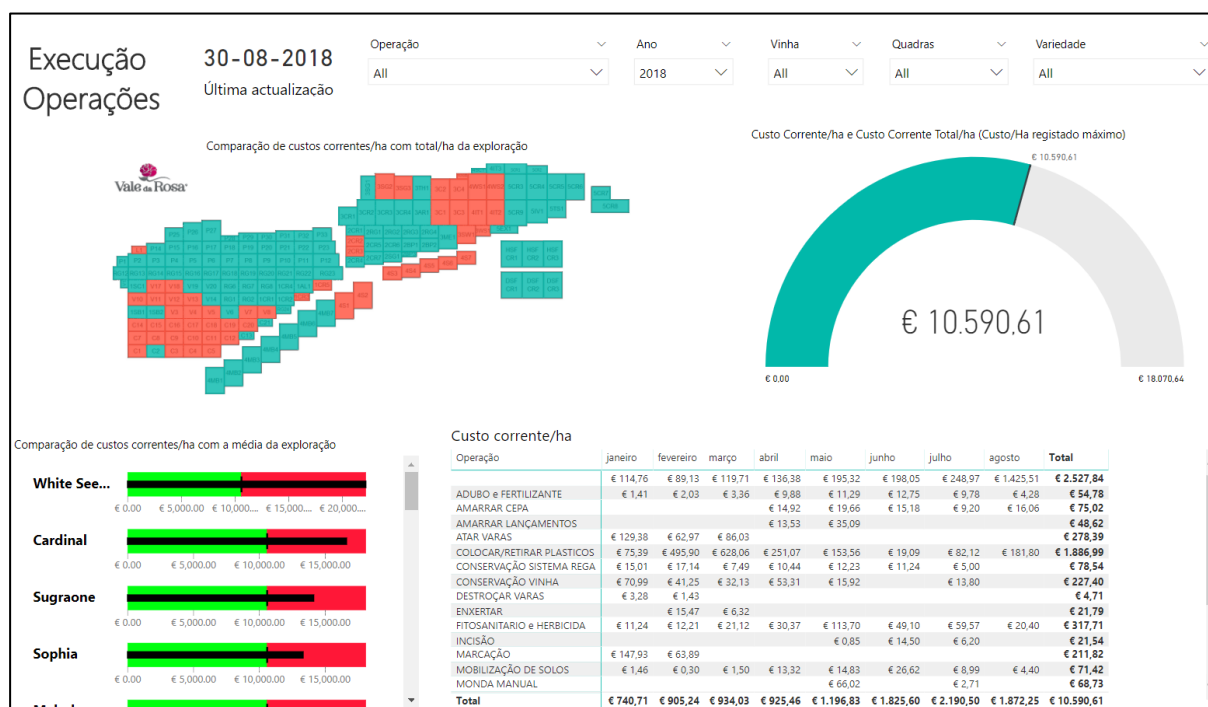


Figura 34 - Painel de execução de operações

De forma a exibir a informação em tempo real, de uma forma intuitiva para o utilizador, este painel utiliza 5 elementos principais, todos eles interligados e com possibilidade de filtragem de todo o painel, ou seja, em função da informação pretendida, o utilizador pode interagir com qualquer elemento do painel de forma a obter a informação que necessita. Em concreto, os cinco elementos são:

- Na extremidade superior, os filtros gerais que possibilitam a seleção célere da informação a ser exibida no painel;
- No canto superior esquerdo, um diagrama representativo de toda a exploração da empresa. Este diagrama é dinâmico e permite a seleção de quadras a partir da própria imagem. Nesta Figura 33, o diagrama encontra-se a expor a comparação do custo operacional das quadras em comparação com o custo global da exploração. A cor encarnada encontram-se as quadras com o custo operacional superior ao custo operacional e a cor azul encontram-se as quadras com o custo operacional inferior;

- No canto superior direito, um gráfico que exhibe o custo operacional da seleção, em função do máximo registado anualmente (relativo à seleção). Apresenta também uma barra que corresponde ao custo operacional da exploração, no entanto visto que não existe seleção nesta figura, a barra coincide com o valor disposto no centro. Este gráfico permite contextualizar as operações executadas em função do registo histórico e em função da situação global da empresa, em tempo real;
- No canto inferior esquerdo, encontra-se um gráfico que ilustra as variedades em função do seu custo operacional e do custo operacional global. Este gráfico pretende demonstrar as variedades em que mais recursos foram utilizados, de forma a poderem ser seleccionadas e posteriormente analisados os diferentes tipos de custos;
- No canto inferior direito, uma tabela que discrimina os custos operacionais em função da operação e da tarefa executada. Esta tabela permite ao utilizador identificar especificamente as operações e tarefas em função dos recursos utilizados.

No último painel do modelo de operações (Figura 34) foi elaborada a exposição da comparação dos objetivos propostos e da execução em 2018.

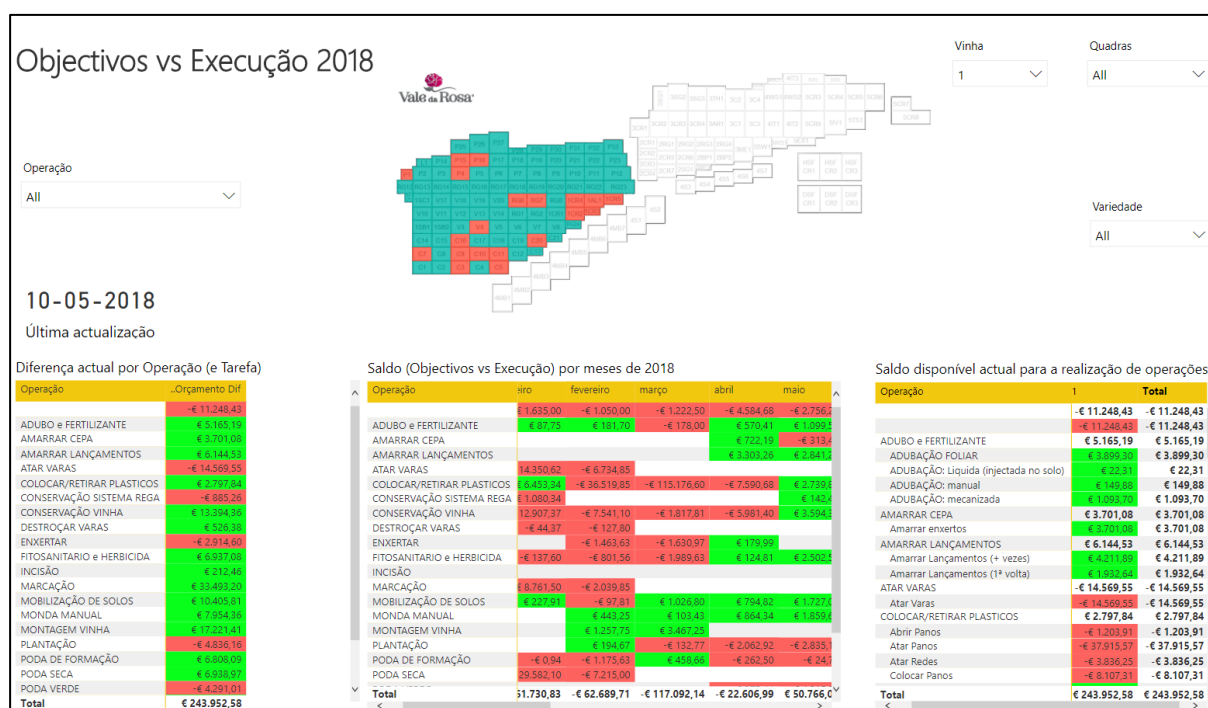


Figura 35 - Painel de diferença entre objetivos e execução de operações.

Neste painel encontra-se seleccionada apenas a vinha 1, como ilustrado graficamente e descrito no filtro “Vinha”. Existem três elementos principais que permitem ao utilizador a análise da situação atualizada da exploração, ao nível das operações. Em concreto esses elementos são:

- Os filtros na forma de lista também presentes nos painéis anteriores, que permitem a seleção de atributos relevantes para a interpretação do painel;

- O diagrama da exploração, que ilustra a situação de cada quadra, sendo que a cor encarnada significa que a execução tem valores superiores aos previstos nos objetivos e a cor verde significa o contrário. Uma vez mais este diagrama está relacionado com o restante painel, permitindo a seleção de filtros a partir do mesmo;
- As tabelas na extremidade inferior, que calculam a diferença entre os objetivos e a execução de operações. A cor verde encontram-se os valores em que o custo execução não ultrapassou o objetivo e, a cor encarnada, o contrário. Estes valores são apresentados em função da totalidade da exploração ou da seleção (tabela esquerda), em função dos meses do ano atual (tabela central) e em função da vinha ou níveis de granularidade inferiores (tabela direita).

## 5. Discussão dos resultados

### 5.1. Avaliação qualitativa dos modelos

Neste capítulo é pretendida a avaliação qualitativa da utilidade do sistema modelado e instalado nas necessidades reais dos diferentes funcionários da Herdade Vale da Rosa. Esta avaliação visa a identificação das limitações e mais valias do sistema segundo aspetos determinantes para o seu correto funcionamento. Para este fim, foi realizada uma reunião com dois elementos da empresa com necessidades de informação distintas e com conhecimento profundo dos diferentes processos produtivos, concretamente, o diretor geral da Herdade Vale da Rosa, Joaquim Praxedes e o responsável pelas operações, Pedro Ramalho.

A avaliação dos modelos criados passa por três pontos distintos, descritos por Shim et al. (2002) como fundamentais para um correto funcionamento de um DSS: (1) a base de dados, (2) a capacidade de processamento de informação do programa e (3) a interface do utilizador. Por outro lado, a avaliação visa também a análise do desempenho dos modelos enquanto DSS que, segundo McBratney et al. (2005), incorporam funcionalidades cruciais para a prática de agricultura de precisão. Finalmente, é pretendida uma análise aos modelos criados enquanto ferramentas para ultrapassar os desafios do tratamento de *Big Data* agrícola na atualidade e nas necessidades futuras da empresa, descritos por Wolfert (2017).

Nesta reunião foram apresentados os modelos criados e, de seguida, postas questões sobre os pontos enunciados (Anexo 1). As questões foram postas simultaneamente aos participantes, de forma a ser realizada uma discussão sobre as respostas de cada um.

#### 5.1.1. Funcionamento do sistema segundo Shim et al. (2002)

A avaliação, neste capítulo, diz respeito aos três aspetos fundamentais para o funcionamento prático de DSS's.

Quanto à base de dados existente na empresa, esta apresenta uma representatividade satisfatória para os intervenientes da reunião. Isto significa que, nas áreas abordadas neste trabalho, os dados não representaram um fator limitante para os modelos criados. Estes dados apresentam um detalhe, volume e velocidade de atualização ajustados às necessidades dos intervenientes da reunião. No entanto, a adição de dados externos à empresa pareceu ser uma opção interessante para o ajuste de tomadas de decisão em função dos mercados em que os seus produtos são escoados. Dados como preços de mercado, *benchmarking* de operações ou equipamentos e informação sobre os consumidores foram citados durante esta avaliação. Foi referido um exemplo de integração no momento da colheita da valorização dos mercados sobre o grau Brix. Isto permitiria à empresa tomar uma decisão mais informada sobre a data de colheita e sobre onde executar a mesma.

Quanto à capacidade de processamento do sistema, as necessidades variam conforme a área de tomada de decisão. Geralmente, uma atualização diária de informação parece ser suficiente para a vasta maioria das tomadas de decisão.

No entanto, foi referida uma situação na época de colheita em que a atualização de duas ocasiões por dia seria interessante. Nesta situação seria útil uma atualização mais frequente, uma vez que equipas de colheita mais experientes demonstram flexibilidade suficiente para alterar os seus objetivos de colheita a meio do dia, sem que com isso seja perdido tempo de mão-de-obra considerável. Neste exemplo, uma atualização de informação de colheita pela manhã e pela tarde, seria bastante relevante para aumentar a eficiência dos trabalhadores na operação em questão.

Quanto à interface do utilizador do sistema, nos modelos finais não pareceu existir falta de informação para as tomadas de decisão em que os painéis incidem, mesmo depois dos dados terem sido alvo de um processamento considerável. Os painéis pareceram intuitivos durante a sua utilização e o seu dinamismo logrou a resposta a diferentes tipos de perguntas que os tomadores de decisão podem necessitar ao longo da campanha, sendo referido que o sistema é prático, tendo em conta a variedade de dispositivos em que pode ser acedido e também a facilidade de utilização.

No entanto, ficaram notas de aspetos a melhorar no sistema, como o direcionamento dos indicadores e dos painéis para a gestão de topo da empresa e a necessidade de existência de um único painel em que a dinâmica seja ainda mais desenvolvida, de forma a que nesse painel seja possível aceder a todos os departamentos da empresa e que sejam apresentados indicadores relevantes para cada uma dessas consultas. Esta sugestão tem como finalidade a instalação de um ecrã tátil num escritório, em que, com pouca interação com o mesmo, seja possível retirar informação do estado atual da empresa.

#### **5.1.2. Desempenho em agricultura de precisão segundo McBratney et al. (2005)**

Para a avaliação do desempenho do sistema criado numa situação de agricultura de precisão foram realizadas questões que têm por base as funcionalidades enunciadas por McBratney et al. (2005).

Da resposta à questão 2.1.1. (Anexo 1) retirou-se que as tecnologias criadoras de dados instaladas na exploração podem ser utilizadas neste sistema, sendo que nem todas o foram, uma vez que as necessidades abordadas neste trabalho não incluem todos os departamentos na empresa. No entanto, o aproveitamento das fontes de informação realmente utilizadas foi bastante eficaz, suprimindo assim a necessidade de consulta direta e posterior tratamento dos dados gerados pelos equipamentos.

Concretamente, neste trabalho foram utilizados e relacionados dados originários das áreas técnica agronómica, de operações, assim como dados físicos de cariz transversal à exploração. Departamentos da empresa como o departamento comercial, de qualidade e de *marketing* ficaram por utilizar. No entanto, o trabalho realizado demonstrou que, no futuro, existe a possibilidade dos

departamentos que não foram abordados serem integrados no sistema BI criado, possibilitando uma tomada de decisão agrícola integrada nas condições da empresa como uma única entidade.

Por outro lado, este sistema de BI permite ao utilizador tomar decisões localmente distintas, possibilitando assim um tratamento diferenciado das diferentes parcelas em que as operações ocorrem. Deste modo, a eficiência de alocação de recursos em função das necessidades localmente específicas aumenta consideravelmente.

### **5.1.3. Processamento de *Big Data* agrícola segundo Wolfert (2017)**

O objetivo deste grupo de questões foi a avaliação do processamento de *Big Data* produzida na Herdade Vale da Rosa pelo sistema criado.

Realizada a questão 3.1.1. (Anexo 1) sobre o valor dos dados criados e a salvaguarda existente ou a criar dos mesmos, as respostas foram que na atualidade já existem planos para proteger os dados na empresa, sendo que já estão implementadas algumas medidas de segurança dos mesmos. No entanto, o valor acrescentado aos dados na forma de processamento é onde a maior valorização se encontra e também onde a proteção é maior.

Os painéis criados neste trabalho possuem como utilizadores finais um pequeno grupo de membros da empresa, associados à direção da mesma. No entanto, numa eventual expansão do sistema para a maior parte dos membros da empresa, a informação disponível para cada um deverá ser controlada de forma a salvaguardar o seu valor. O objetivo futuro não será o livre acesso a toda a informação por parte dos diferentes membros, mas sim o acesso personalizado em função das necessidades do utilizador para cumprir as suas tarefas.

Quanto aos erros associados à emissão de dados, foi reconhecido que estes existem e representam um desafio para a empresa. Um dos aspetos debatidos foi a criação de algoritmos (no sistema BI criado) que identificasse dados possivelmente errados e que notificasse o utilizador. Desta forma a segurança da tomada de decisão baseada no sistema BI aumentaria. De outra forma sem ser a descrita, a identificação de erros pelos meios anteriormente utilizados neste regime de criação de *Big Data* corresponderia a um custo de manutenção da base de dados bastante avultados, sendo um processo pouco eficiente e demorado.

Quanto à integração de *Big Data* no modelo de negócios da empresa, esta será sempre um processo gradual, em que a utilização de *Big Data* de forma definitiva terá de ser comprovada ao longo da sua integração nos diferentes departamentos da empresa. Sendo que existe uma grande variabilidade de necessidades de informação, a integração de sistemas de BI será obrigatoriamente precedida de um estudo das necessidades dos diferentes utilizadores, de forma a garantir que a informação que é utilizada por cada um é realmente relevante para os seus processos.

## 5.2. Análise dos resultados da modelação e implementação do sistema

O aspeto prático e inovador deste trabalho consistiu fundamentalmente nas fases de modelação e implementação, sendo que, durante as mesmas, foram sendo identificadas as necessidades de informação nas áreas trabalhadas e também a forma como os dados produzidos podem ser aproveitados de uma forma mais eficiente.

Como referido por Wolfert (2017), “o *Big Data* sem programas analíticos de BI não representa valor”. Neste trabalho o objetivo do sistema desenvolvido foi a adição de valor acrescentado aos dados produzidos. Neste contexto foram obtidos os resultados que se descrevem de seguida.

Por um lado, os painéis criados auxiliam o processo de tomada de decisão dos seus utilizadores, podendo esta ser executada somente através da consulta dos painéis. A qualidade da tomada de decisão não é proporcional à quantidade de informação, mas sim à qualidade de informação (Saaty, 2008), e os modelos lograram reduzir a quantidade de informação disponível para o utilizador para aumentar a qualidade da mesma. A tomada de decisão consiste em conhecer o problema, avaliar a importância de uma decisão, incorporar os critérios em que a tomada de decisão assenta, as suas consequências e comparar ações alternativas (Shim et al., 2002), e os resultados obtidos desde o modelo de demonstração inicial até aos modelos finais revelam que, para este processo, os modelos BI revelam a capacidade de auxiliar em todas estas fases.

Por outro lado, um resultado obtido foi a criação de painéis inteligíveis pelo seu utilizador final e dinâmicos, de forma a alterarem as suas informações de acordo com o pretendido em cada momento. Como referido por Lindblom et al. (2017), a limitação da utilização da informação produzida não está do lado dos consumidores, mas sim dos DSS's não lograrem uma exposição simples e representativa da informação obtida. No caso deste trabalho, uma das prioridades foi exatamente a identificação das necessidades dos utilizadores e só posteriormente a elaboração dos modelos. Desta forma não foi o DSS a indicar o que o agricultor precisa de conhecer, mas sim o agricultor a “indicar” as informações que o DSS necessitava de expor. Em suma, este foi um resultado de importância relevante, confirmado pela avaliação qualitativa do modelo, pois representava uma das maiores limitações para a implementação do sistema, uma vez que, segundo Wolfert (2017), para um correto funcionamento da utilização de *Big Data* é necessário o conhecimento das particularidades dos processos organizacionais para a introdução integrada da utilização de *Big Data* nesses mesmos processos.

Um outro resultado obtido neste trabalho diz respeito à automatização do funcionamento do sistema. Embora a ligação em tempo real aos servidores da Herdade Vale da Rosa não tenha sido realizada, este sistema está programado para uma atualização simplificada dos dados através da introdução dos mesmos numa pasta predefinida. Tendo em consideração que um dos três V's que define *Big Data* é a substancial velocidade de criação de novos dados (Laney, 2001), a forma como os mesmos são introduzidos no sistema teria obrigatoriamente de ser automatizada ou bastante simplificada. Neste trabalho, o resultado, não tendo sido o ideal, que seria a ligação aos servidores da empresa com a finalidade da atualização, foi, no entanto, o mais aproximado possível desse mesmo resultado ideal. O programa criado tem a capacidade de receber os dados não processados numa pasta e, automaticamente, processar os mesmos e atualizar os painéis num curto período.



Durante a elaboração deste trabalho foram igualmente obtidos resultados na forma de limitações, tais como as previstas nos trabalhos de Sonka (2016) e Wolfert (2017). O principal desafio a ultrapassar, aquando da modelação e automatização, foi o formato dos dados originais. Um exemplo desta limitação são os dados do programa inicial de consumos (Figura 5 e Figura 6), que se encontram num formato não uniforme, em que os ficheiros possuem diferentes tabelas para cada variedade, enquanto o formato ótimo para o processamento automatizado dos dados no Power BI® consiste numa única tabela (Tabela 2) com vários atributos. Neste caso a limitação teria de ser ultrapassada pela alteração da emissão dos dados para um formato idêntico e uniforme entre emissões.

No exemplo enunciado e noutras situações semelhantes, o trabalho analítico e a automatização do mesmo encontram-se limitados pela complexidade e heterogeneidade do formato original.

“A analítica de dados é o maior fator diferenciador de *Big Data*” (Sonka, 2016), o que significa que qualquer limitação ao tratamento analítico e à sua automatização constitui uma severa perda de valor dos dados produzidos.

### 5.3. Perspetivas

Posteriormente à apresentação final dos modelos em funcionamento, a empresa Herdade Vale da Rosa mostrou-se interessada em continuar o trabalho de processamento de *Big Data* e reconheceu que no futuro deverá integrar esta tecnologia em mais departamentos da empresa, para obter uma visão holística dos seus processos, mas também lograr a mais profunda integração dos mesmos entre si. Para tal, baseado nos resultados deste caso de estudo, futuramente devem ser levados em consideração os seguintes pontos:

- **Reformulação dos dados originais para otimização das ferramentas BI**

No momento de introdução de processos e tecnologias criadoras de dados não existia a ideia de implementação de um sistema de BI e, por conseguinte, a maior parte dos dados produzidos não se encontram no formato mais indicado para o seu processamento integrado com os outros dados da empresa. Esta situação é particularmente notória nos dados de programa inicial de consumos (Figura 5 e Figura 6), assim como os dados relativos à evolução fenológica das videiras (Figura 11).

No futuro, para um funcionamento otimizado do sistema de BI é necessária uma reestruturação da emissão de dados, pois embora os modelos criados consigam utilizar a informação atual, o processo de preparação é pouco eficiente e existem dados que se perdem, assim como erros resultantes da heterogeneidade no formato e conteúdo entre emissões, existindo assim um prejuízo na forma de valor dos dados perdidos.

- **Conexão automatizada *online* em tempo real**

Os dados existentes na Herdade Vale da Rosa encontram-se em constante atualização. Uma das mais valias do sistema implementado é a capacidade de atualizar os dados em que se baseia em tempo real. Para a maior parte dos processos de tomada de decisão, a informação disponível nos

modelos representa um maior valor quanto mais recentemente tiver sido emitida. Para alcançar a atualização em tempo real, o programa de BI necessita de estar conectado ao servidor da empresa, assim como a todas as fontes de dados (automatizadas ou não). Desta forma, as decisões irão ser realizadas com base em informação atualizada até ao momento da tomada da mesma.

- **Perfis**

As necessidades de informação de cada funcionário da empresa são variáveis consoante a sua função, significando isto que os painéis para cada funcionário deverão ser, necessariamente, distintos.

Para atingir o objetivo de painéis personalizados em função do seu utilizador, no futuro deverão ser traçados perfis de utilizador, ou seja, o criador do sistema necessitará de conhecer os processos pelos quais o funcionário é responsável e a informação que o mesmo necessita durante a execução das suas funções. Somente desta forma o sistema BI poderá ser útil e utilizado por diferentes funcionários da empresa.

Por outro lado, a criação de perfis permitirá uma maior eficiência na gestão da segurança dos dados, sendo que cada utilizador só terá acesso à informação que necessita para cumprir as suas funções.

- **Diferentes áreas de aplicação**

Por fim, o sistema aplicado neste trabalho às operações culturais e aos consumos de fertilização pode ser aplicada noutras áreas já existentes na empresa. A área comercial seria um exemplo, uma vez que os dados existentes na empresa também já se encontram em formato digital e seria uma boa oportunidade para integrar dados externos disponíveis, como indicadores de mercado disponibilizados publicamente e gerar novas métricas de negócio.

No entanto, não só a área comercial poderá ser integrada neste sistema, a área agrícola dispõe de mais informação informatizada, mas também a área de recursos humanos, marketing e financeira poderão vir a ser integradas neste sistema.

Uma vez realizada a integração de várias áreas no sistema, existe uma grande oportunidade de relacionamento de informação entre os diferentes departamentos, o que pode vir a revelar uma grande mais valia para o funcionamento da empresa como um só corpo (Sonka, 2016).

## 6. Considerações finais

Com a realização deste trabalho verificou-se que foi possível a criação e implementação prática de um programa de processamento de *Big Data* em agricultura, melhorando processos de tomada de decisão. Desta forma, os objetivos apresentados neste estudo foram cumpridos, procurando primeiramente realizar a análise dos dados disponíveis na empresa, o seu formato e a forma em que se encontravam disponíveis e, posteriormente conhecer as necessidades dos responsáveis dos processos em estudo, no que respeitava às suas tomadas de decisão, para ser possível a implementação de modelos que fossem utilizados regularmente. Finalmente, foi realizada uma avaliação qualitativa do sistema instalado, que revelou o interesse na continuação da implementação deste sistema em outras áreas da empresa, cumprindo assim o terceiro e último objetivo proposto.

Neste sentido, foi construído um modelo que permitiu não só comprovar que os sistemas de BI possuem um lugar importante nas empresas agrícolas, ao melhorar a eficiência dos seus processos, mas também que o conhecimento técnico agrícola aliado ao conhecimento informático representa uma mais valia na implementação destes mesmos sistemas. Tendo ainda a exatidão e fiabilidade do sistema aos mais distintos níveis de granularidade de análise sido um ponto central de toda a metodologia realizada.

A sustentabilidade e, em particular, a intensificação sustentável dos processos agrícolas representa um desafio global emergente (Garnett et al., 2013). O contributo deste sistema de BI para ultrapassar o mesmo, é também outro aspeto importante que se pode retirar deste trabalho, na medida em que, ao expor informações detalhadas e atualizadas, os responsáveis pelos processos agrícolas serão capazes de tomar decisões de uma forma mais eficiente no que diz respeito à utilização de insumos. Serão assim capazes de produzir mais com menos, ou seja, produzir mais, pois é possível uma identificação das necessidades das plantas a uma escala bastante reduzida, e utilizar menos, pois a alocação de recursos só será realizada quando e onde existir necessidade para tal, num curto tempo de resposta, graças à quase constante atualização das informações disponíveis.

Em suma, com este trabalho é possível concluir que o processamento de dados produzidos, tendo por base conhecimento agronómico, representa uma oportunidade para uma agricultura mais eficiente na utilização de recursos, mas também abre a oportunidade para uma melhor produção e otimização da produtividade. Neste aspeto, é importante referir que o tomador de decisão será sempre o responsável por esta intensificação sustentável. No entanto, as informações nas quais se baseia limitam inevitavelmente a qualidade da sua decisão, aspeto que reforça a importância do sistema de auxílio à tomada de decisão desenvolvido neste trabalho.



## 7. Referências bibliográficas

Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computers and electronics in agriculture*, 44(1), 71-91.

Auernhammer, H. (2001). Precision farming—the environmental challenge. *Computers and electronics in agriculture*, 30(1-3), 31-43.

Bellon-Maurel, V., Peters, G. M., Clermidy, S., Frizarin, G., Sinfort, C., Ojeda, H., ... & Short, M. D. (2015). Streamlining life cycle inventory data generation in agriculture using traceability data and information and communication technologies—part II: application to viticulture. *Journal of cleaner production*, 87, 119-129.

Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). Precision agriculture and sustainability. *Precision agriculture*, 5(4), 359-387.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5), 662-679.

Braga, R. (2009). Inovação e tecnologia na formação agrícola. Manual de viticultura de precisão.

Bronson, K., & Knezevic, I. (2016). Big Data in food and agriculture. *Big data & society*, 3(1), 2053951716648174.

Chen, Y., Yu, J., & Khan, S. (2010). Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. *Environmental modelling & software*, 25(12), 1582-1591.

Cooper, M., & Mell, P. (2012). Tackling big data. Federal computer security managers' forum (pp. 728-729).

Crookston, R. K. (2006). A top 10 list of developments and issues impacting crop management and ecology during the past 50 years. *Crop science*, 46(5), 2253-2262.

Elbashir, M. Z., Collier, P. A., & Davern, M. J. (2008). Measuring the effects of business intelligence systems: The relationship between business process and organizational performance. *International journal of accounting information systems*, 9(3), 135-153.

Elbashir, M., & Williams, S. (2007). BI impact: The assimilation of business intelligence into core business process. *Business intelligence journal*, 12(4), 45-54.

Erdal, G., Esengün, K., Erdal, H., & Gündüz, O. (2007). Energy use and economical analysis of sugar beet production in Tokat province of Turkey. *Energy*, 32(1), 35-41.

FAO (1996) Declaration on world food security. World Food Summit, FAO, Rome

Feder, G., Just, R. E., & Zilberman, D. (1985). Adoption of agricultural innovations in developing countries: A survey. *Economic development and cultural change*, 33(2), 255-298.

Fountas, S., Wulfsohn, D., Blackmore, B. S., Jacobsen, H. L., & Pedersen, S. M. (2006). A model of decision-making and information flows for information-intensive agriculture. *Agricultural systems*, 87(2), 192-210.

Garnett, T., Appleby, M. C., Balmford, A., Bateman, I. J., Benton, T. G., Bloomer, P., ... & Herrero, M. (2013). Sustainable intensification in agriculture: premises and policies. *Science*, 341(6141), 33-34.

Gitelson, A. A., Vina, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., & Leavitt, B. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophysical research letters*, 30(5).

Gray, J., & Rumpe, B. (2015). Models for digitalization. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0109564105001314> (acedido a 1 de outubro de 2018).

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future generation computer systems*, 29(7), 1645-1660.

Hokazono, S., & Hayashi, K. (2012). Variability in environmental impacts during conversion from conventional to organic farming: a comparison among three rice production systems in Japan. *Journal of cleaner production*, 28, 101-112.

Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in human geography*, 3(3), 262-267.

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & society*, 3(1), 2053951716631130.

Laney, Doug. "3D data management: Controlling data volume, velocity and variety." META group research note 6.70 (2001): 1. Disponível em: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (acedido a 30 de julho de 2018).

Lindblom, J., Lundström, C., Ljung, M., & Jonsson, A. (2017). Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies. *Precision agriculture*, 18(3), 309-331.

Mamo, M., Malzer, G. L., Mulla, D. J., Huggins, D. R., & Strock, J. (2003). Spatial and temporal variation in economically optimum nitrogen rate for corn. *Agronomy journal*, 95(4), 958-964.

McBratney, A., Whelan, B., Ancev, T., & Bouma, J. (2005). Future directions of precision agriculture. *Precision agriculture*, 6(1), 7-23.

Mulla, D. J., & Schepers, J. S. (1997). Key processes and properties for site-specific soil and crop management. In *The state of site specific management for agriculture*. Amer. Soc. Agronomy.

- Olszak, C. M., & Ziemba, E. (2007). Approach to building and implementing business intelligence systems. *Interdisciplinary Journal of Information. Knowledge, and management*, 2(1), 135-148.
- Pinstrup-Andersen, P. (2009). Food security: definition and measurement. *Food security*, 1(1), 5-7.
- Reinschmidt, J., & Francoise, A. (2000). Business intelligence certification guide. IBM International technical support organisation.
- Robert, P. (1993). Characterization of soil conditions at the field level for soil specific management. *Geoderma*, 60(1-4), 57-72.
- Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International journal of services sciences*, 1(1), 83-98.
- Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision support systems*, 33(2), 111-126.
- Sonka, S. (2016). Big data: fueling the next evolution of agricultural innovation. *Journal of innovation management*, 4(1), 114-136.
- Srbinska, M., Gavrovski, C., Dimcev, V., Krkoleva, A., & Borozan, V. (2015). Environmental parameters monitoring in precision agriculture using wireless sensor networks. *Journal of cleaner production*, 88, 297-307.
- Stafford, J. V. (2000). Implementing precision agriculture in the 21st century. *Journal of agricultural engineering research*, 76(3), 267-275.
- Stern, N. (2008). The economics of climate change. *American Economic Review*, 98(2), 1-37.
- Suter, C.R. (1992). The professional farm manager and management Consultant. Retus, Inc., West
- van Meensel, J., Lauwers, L., Kempen, I., Dessein, J., & Van Huylenbroeck, G. (2012). Effect of a participatory approach on the successful development of agricultural decision support systems: the case of Pigs2win. *Decision support systems*, 54(1), 164-172.
- Verdouw, C. N., Beulens, A. J., Reijers, H. A., & van der Vorst, J. G. (2015). A control model for object virtualization in supply chain management. *Computers in industry*, 68, 116-131.
- Wang, J. J., Jing, Y. Y., Zhang, C. F., & Zhao, J. H. (2009). Review on multi-criteria decision analysis aid in sustainable energy decision-making. *Renewable and sustainable energy reviews*, 13(9), 2263-2278.
- Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming—a review. *Agricultural systems*, 153, 69-80.

Wolfert, S., Sorensen, C. A. G., & Goense, D. (2014). A future internet collaboration platform for safe and healthy food from farm to fork. In Global Conference (SRII), 2014 Annual SRII(pp. 266-273). IEEE.

Ye, J., Chen, B., Liu, Q., & Fang, Y. (2013). A precision agriculture management system based on Internet of Things and WebGIS. In Geoinformatics (GEOINFORMATICS), 2013 21st International Conference on (pp. 1-5). IEEE.

Yeoh, W., & Koronios, A. (2010). Critical success factors for business intelligence systems. *Journal of computer information systems*, 50(3), 23-32.

Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision agriculture*, 13(6), 693-712.



## Anexos



## **Anexo 1 – Documento de questões realizadas na fase de avaliação dos modelos criados**

### **Avaliação qualitativa do sistema BI**

#### **Grupo de questões 1 – Baseadas em Shim et al. (2002)**

##### **Base de dados**

**1.1.1.** A base de dados é suficientemente completa para responder às necessidades de tomada de decisão? O que poderia ser acrescentado à mesma?

**1.1.2.** Ajudaria a tomada de decisão a adição de dados externos, como preços médios de mercado regionais, nacionais ou internacionais?

Para um apoio à decisão representativo, a ferramenta terá de dispor de uma **base de dados representativa**, bem como de um grau de especificidade suficiente para responder ao grau de especificidade da solução pretendida. Esta base de dados deverá ter acesso a dados internos e externos correspondentes ao problema a abordar, a informação sobre o problema e as capacidades de resolução do mesmo, e acesso a conhecimento de forma a tratar os dados e as informações disponibilizadas (Shim et al., 2002).

##### **Capacidade de processamento**

**1.2.1.** O processamento de informação (velocidade e tipo de indicadores produzidos) da ferramenta apresentada é adequado às necessidades da empresa?

A capacidade de processamento é um dos aspetos principais de um DSS, uma vez que o objetivo do mesmo é o relacionamento de um grande volume de dados, de uma forma que não seria possível ser realizada por um indivíduo, num espaço de tempo útil (Shim et al., 2002).

##### **Interface do utilizador (IU)**

**1.3.1.** A IU é inteligível para a interpretação da situação da empresa? Existe algum tipo de informação que parece ser negligenciado para a exibição simples da informação?

**1.3.2.** A interatividade dos painéis é um aspeto importante aquando da análise da informação? O processo de adaptação a esta forma de apresentação é uma mais valia ou é redundante com as análises que seriam possíveis realizar em formatos mais tradicionais (tabelas simples)?

**1.3.3.** Os painéis logram conferir uma visão holística da situação atual da empresa?

Os resultados de um DSS devem ser apresentados numa interface de utilizador inteligível, o que significa que deve lograr a interpretação do seu utilizador, sem que com isso perca a representatividade dos seus resultados, através de gráficos, tabelas e interatividade entre o DSS e o utilizador (Shim et al., 2002).

#### **Grupo de questões 2 – Baseadas em McBratney et al. (2005)**

**2.1.1.** As tecnologias instaladas na exploração agrícola estão/podem ser representadas neste sistema? Ou estão as informações conferidas pelas mesma subaproveitadas?

**2.1.2.** Permite este sistema tomar decisões localmente distintas, com um grau de especificidade suficiente para as diferentes áreas?

**2.1.3.** Dentro das áreas da empresa abordadas, existe alguma fonte de informação que poderia ser aproveitada para este processamento e não o é?

Essas ferramentas devem contemplar funções como **a utilização e integração de novas tecnologias instalados na exploração**, realizar análises de custo/benefício comercial em áreas específicas da parcela onde está instalada a cultura, **integrar tecnologias de forma a obter desde uma visão holística da exploração, a uma especificidade local cultural** e, finalmente, **o estabelecimento de protocolos distintos para uma abordagem de gestão localmente específicas em diferentes culturas**.

Em suma, as ferramentas de apoio à tomada de decisão não deverão ser somente ferramentas de indicadores culturais, mas sim englobar as características dos processos a realizar e incorporar a dimensão económica na altura de prestar o apoio à tomada de decisão (McBratney et al., 2005).

#### **Grupo de questões 3 – Baseadas em Wolfert (2017)**

**3.1.1.** Tendo em consideração o valor crescente dos dados produzidos, existe preocupação por parte da empresa nos aspetos de privacidade e segurança? O facto de os dados se encontrarem *online* representam alguma preocupação acrescida?

**3.1.2.** Existem erros associados à emissão de dados? Esses erros representam fatores limitantes no momento da tomada de decisão? Será necessária a implementação de mecanismos que identifiquem dados errados?

**3.1.3.** A criação de modelos analíticos por parte de um agrónomo em detrimento de um técnico analítico confere um benefício para a empresa? Quais as preocupações e as mais-valias numa decisão deste tipo?

**3.1.4.** Numa empresa com as características da Herdade Vale da Rosa, a integração de *Big Data* no modelo de negócios da mesma assume alguma importância? Num exercício especulativo, quais os principais riscos que existiriam numa integração máxima de *Big Data* em todos os processos de tomada de decisão na empresa?

O surgimento de possibilidades de utilização e processamento de *Big Data* no contexto de produção agrícola levanta algumas limitações presentes e futuras que representarão desafios a ultrapassar (Wolfert, 2017).

Limitações relacionadas com a **propriedade dos dados produzidos, a sua privacidade e segurança**, em que embora já se encontrem a ser legislados e regulados alguns aspetos no contexto nacional e internacional, ainda é necessário maior salvaguarda dos interesses dos proprietários dessa informação, sem com isso abrandar os desenvolvimentos tecnológicos que permitem o avanço da valorização da informação (Wolfert, 2017).

Um outro desafio da utilização de *Big Data* abrange a qualidade dos dados obtidos e a sua representatividade, uma vez que ao ser tratado um elevado volume de registos (a uma elevada velocidade e com variedade nos dados obtidos) a tendência segue para o **aparecimento de dados incorretos dentro das bases de dados**, o que obrigará à criação de mecanismos de controle de dados adaptados aos contextos específicos de cada organização, não existindo para já uma solução universal para todas as organizações, uma vez que os contextos produtivos se alteram em cada situação empresarial particular (Wolfert, 2017).

Em seguimento do desafio anterior, o processamento analítico eficiente e adequado a cada organização requer a **junção do conhecimento de ciência analítica, ao conhecimento da área em que a organização opera**. Isto é um fator bastante limitante para qualquer base de dados eficiente no contexto organizacional onde se encontra. Esta limitação é relevante e representa um desafio, uma vez que para uma base de dados ser otimamente construída e posteriormente processada é necessária a junção de duas áreas: **uma transversal às organizações e outra específica à atividade da organização**. Somente desta forma se conseguirá um processamento de *Big Data* eficiente e adequada às necessidades de cada organização em particular. Para ultrapassar este desafio será necessário uma relação inteligente **entre um especialista em ciência analítica e um especialista da área operacional da organização** (Wolfert, 2017).

**A integração de *Big Data* no modelo de negócios de uma empresa** necessita de ser realizada de uma forma sustentável, o que representa um outro desafio, uma vez que por um lado, o modelo de negócios da organização beneficia com o máximo de informação relevante processada para os seus processos produtivos mas, por outro lado, uma transição demasiado célere para uma dependência do processamento de *Big Data* causará um risco em que alguns aspetos, anteriormente não contemplados, aquando da integração da utilização de *Big Data*, possam vir a revelar fragilidades na forma de processamento desta informação. Torna-se necessário um estudo extensivo das particularidades dos processos organizacionais para a introdução integrada da utilização de *Big Data* nesses mesmos processos (Wolfert, 2017).

Finalmente, outro desafio na utilização de *Big Data* será o desenvolvimento de plataformas abertas para o desenvolvimento de soluções e inovações no que diz respeito a diferentes formas de processamento de dados com as características da *Big Data* – elevados volume, velocidade e variedade (Laney, 2001) – de forma a acelerar inovações, mas também para conferir ao setor primário uma posição mais consolidada na cadeia de valor (Wolfert, 2017).